CENTRE FOR BUSINESS,
INFORMATION TECHNOLOGY AND ENTERPRISE

Thesis – INFO902

‘Samthripthi’ - A New Product Rating Model Based on Customer Reviews

Submitted by

Sanjo Vincent - 19488551

STATEMENT OF AUTHENTICITY

By submitting this work, I declare that it is entirely my work, except the sections correctly described and cited in my submission. It follows any word restrictions that have been specified, as well as the specifications and requirements described in the coursework guidelines and any other applicable programme module declaration. By submitting this work, I acknowledge that I have read and agreed to the academic misconduct rules and standards, including those against plagiarism, as outlined in the programme handbook. I also agree that this dissertation will face several intellectual credibility tests.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to the following individuals for their unwavering assistance and help in the execution of this study; I believe that without them, the study's goal would have been impossible to accomplish.

First and foremost, I want to thank Dr Monjur Ahmed (Principal Supervisor) and Dr Diab Abuaiadh (Co-Supervisor) for their unwavering support of my research. Through their thorough assistance and supervision, they both led me through each phase of this research project and helped me stay on track. Both of my supervisors inspired me and suggested positive constructive input on my ideas. Their motivational style encouraged me to work hard on this research.

Dr Guess Wilkinson, as a Team Manager, gave me the chance to do this research. In addition to my Team Manager, I would like to thank my institution, Waikato Institute of Technology (Wintec), Hamilton, for the outstanding assistance provided by all staff members in completing this research successfully.

I can't forget the names Dr Kay Fielden and Dr Reza Rafeh, who assisted me greatly in identifying the research gap and preparing the research proposal. I'd like to express my gratitude to them for their inspiration and assistance in finalising my research proposal.

Last but not least, I would like to thank anyone who assisted me in this study, both directly and indirectly. I am grateful to everyone who helped me financially and morally with this study.

Sanjo Vincent Vadakkechundeveli Ouseph
Student ID – 19488551
Master of Applied Information Technology
Waikato Institute of Technology (Wintec)
Hamilton City Campus, Hamilton, New Zealand

ABSTRACT

Over the past few decades, the shopping culture especially the eCommerce at has large witnessed tremendous growth and revolutionary changes have been the cornerstone for this multi-faceted progress. Nowadays, no country in this world is untouched by the phenomenon of eCommerce. Detection and recognition of products through user reviews is the new trend fad on eCommerce. With the variety of realistic uses in marketing, sales, media, customer preferences, and policy the user review analysis playing a vital role. Currently, one of the recent developments in data science is the study of user feedback and product ranking computation using sentiment analysis. The previous product rating solutions are purely based on sentiment analysis through different models suggested by many researchers. Based on the previous studies and written literature, the researcher determined that there isn't a solution that takes into account things like review sentiment, review time, and review helpfulness. The researcher looked into the topic's validity and concluded that the analysis regarding a new solution for product ranking is relevant when the above-mentioned considerations are taken into account. The research aims to propose a new model on product rating in e-commerce by defining the review sentiment, review relevance (review time) and review helpfulness. This research following a quantitative research methodology using DSR and fine-tuned BERT language model with 0.88 as precision, 0.89 as recall, 0.88 as F1 score and 0.88 as accuracy for textual evaluations of customer reviews extracted from e-commerce websites such as amazon.com. The researcher suggesting five different algorithms to calculate review helpfulness score, review time score, review sentiment score, overall review rating score and product rating score. Then the eCommerce service provider will show the consumer the product's total star rating as well as the number of feedbacks based on the previously mentioned algorithms. The study found that new and positive reviews have a greater influence on the product ranking score than old and unhelpful reviews. Finally, this research found that the review score calculation based on the variables such as review time, review helpfulness along review sentiments will provide a much more reliable product rating in e-commerce platform than the existing methods. As a consequence of this research discovering how much a consumer is happy with a certain product, the researcher named this proposed model ‘Samthripthi' (സംതൃപ്തി), which is a term from the Malayalam language that indicates satisfaction.

Keywords: - BERT model, Customer Reviews, Consumer Satisfaction eWOM, Product Rating, Review Sentiment, Sentiment Analysis

Contents

List of Figures

List of Tables

List of Equations

1. INTRODUCTION

1.1 Introduction

The Internet's latest craze is sentiment detection and recognition. Sentiment analysis has a wide range of practical applications in marketing, sales, media, customer preferences, and policymaking (Burghardt, 2011; O'Connell et al., 2014; Gadekallu et al., 2019). The study of customer feedback and product rating computation, which includes sentiment analysis, is one of the most recent discoveries in data science (Saldaña, 2018; Dominik, 2017). So, the studies of the user reviews are one of the most required solutions in eCommerce to understand the user acceptance of particular products which are dealing by the eCommerce providers. In this chapter, we discuss the overview in section 1.2 following by objectives and contribution section 1.3. Section 1.4 deals with the article structure, and section 1.5 concluding this chapter.

1.2 Overview

The customers' satisfaction influences a company's profit margin, and what defines the happiness of consumers is the nature and quality of the commodity delivered to them (Larson & Denton, 2014; Zainal et al., 2017). Customer satisfaction is considered the significant factor in every company's profit and is considered quality management (Zainal et al., 2017; López & Sicilia, 2014). So, the evaluation of the satisfaction level of customers plays a vital role in the pathway of the existence of the company. In previous decades, before introducing digital shopping, Word of Mouth (WoM) was considered a significant factor in different industries (Vo et al., 2017; Mellinas & Reino, 2019; Ardyan & Sudyasjayanti, 2020). The invention of mobile technologies and digital shopping platforms and e-commerce websites redefined word of mouth and consumers started to share their opinions through different digital platforms. These comments have subjective aspect-based viewpoints that help buyers and sellers get a better understanding of the consumables. This reviewing process replaces the shopping culture, and the customers get more opportunity to hear about the user experiences and views in a particular product or a specific company in detail from different personalities in society (Maduretno & Junaedi, 2021; Xue, 2019). The customers share their experiences using either positive or negative sentences as online reviews, and these are considered critical information to the new customers for their purchase decisions (Gruen et al., 2006; Ngarmwongnoi et al., 2020).

Because of the increasing usage of the Internet and social networking systems, consumers typically compare multiple product options by obtaining public opinions from online reviews before making a purchase decision (Danniswara et al., 2020; Turkyilmaz & Poturak, 2017). Online shopping has increasingly become the mainstream shopping alternative of customers, thanks to the exponential growth of the Internet and the eCommerce industry. Online product reviews from people who have purchased or used the items are useful sources of knowledge when purchasing online. Online shopping has increasingly become the mainstream shopping alternative of customers, thanks to the exponential growth of the Internet and the eCommerce industry (Zhang & Park, 2015; Gruen et al., 2006). Online product reviews (OPRs) from people who have purchased or used the items are useful sources of knowledge when purchasing online (Sasikala & Sheela, 2020). Previous studies such as Liu and Liao (2019) and Yang and Zhu (2018) have suggested that the product review analysis empowers the decision-making and product rating. However, this is a prolonged and work concentrated process. Some websites holding user reviews have a primary product comparison function that compares numerous prospective products in several features to help customers make an informed purchase decision. As per the researcher's view, this type of comparison system often uses an average rating method based on customer ratings for each product, but this system is not efficient or reliable because there are certain limitations, such as: (1) If the customer not purchased the item but they have done its review or rating; (2) the relevancy of the review time - the review time should be relevant to the technology and its launch date; and (3) the average rating is solely based on customer ratings.

1.3 Objective

The majority of reviews and ratings express the quality of a product, but a prospective buyer does not have the time or resources to read all reviews for decision-making purposes. One of the product comparative studies by Zha et al. (2014) relies heavily on consumer reviews rather than other factors. Sometimes the ratings corresponding to the reviews make some contradictions to the readers. For instance, a reviewer reviewed a product as average and marked a 4-star rating will lead the buyer to some thought processes. As a result, an automatic rating system focused on consumer textual reviews is required to assist in this process.

The researcher defines the research objective of this research as follows:

- To suggest an automated product rating system by evaluating customer reviews along with the unique attributes such as review time, sentiment orientation of user review, and the impact of review helpfulness.

This analysis would weigh considerations such as review time, the result of the sentiment evaluation of customer review, and the impact of review helpfulness to incorporate an accurate and consistent rating. Due to the research objective, the researcher calls this proposed model as 'Samthripthi' (സംതൃപ്തി), which is a word from the language Malayalam which means satisfaction.

1.4 Contribution

The main contribution of this research include:

(1) How the factors like review time, review helpfulness and review sentiments affect an automated product rating.

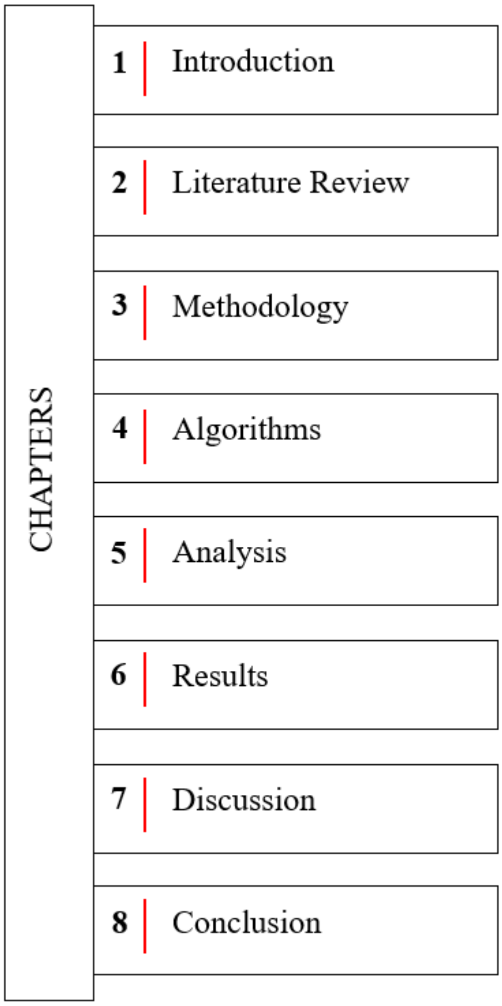
(2) Develop and empirically test a product rating algorithm to show its viability and utility.

Centred on a proposed modern predictive product rating algorithm, this study investigates the impact of variables such as consumer reviews' sentiment analysis results, review time, and the score for the review helpfulness. A structured literature review is provided with a conceptual context to make the idea easier for the study. The details on the importance of WoM, the evolution of sentiment analysis, and previous studies in product rating was gathered through a study of literature. For this study, the researcher developed a model to describe the relationship between independent variables and dependent variables and the hypothesis. This study employs a valid experimental quantitative research method. The new model is developed using user feedback that is freely accessible on e-commerce websites such as Amazon.com, and it is analysing in the comparison section. The new model discusses in the discussion section before being summarised in the conclusion section. The descriptions of the publications used in this writing are included in the reference section.

1.5 Report Structure

The research subject is introduced in the first chapter. Following the overview, the report's second chapter reflects on the literature review, split into many sub-sections. In the third chapter, the research methodology is specified. The collected data is analysed using the proposed model 'Samthripthi' in chapter 4. Chapter 5 discusses the data that has been analysed and the results that have been obtained. The discussions regarding this study have been discussed in the discussion chapter, which is chapter 6, and the research report's conclusion is found in chapter 7. Figure 1.1 illustrating the structure of this report in eight sections.

Figure 1.1 - Report Structure



1.6 Conclusion

According to the researcher, one of the most critical considerations when designing a product rating algorithm is WOM or eWOM visibility, recognition, and impact on purchasing. So, analysing user feedback and other external variables would aid in creating a new product rating algorithm. This would be very beneficial to businesses in obtaining genuine approval and product ratings on their products. The literature review that is used to frame this study is presented in the following section.

2. LITERATURE REVIEW

2.1 Introduction

In this literature review section, the researcher is reviewing a collection of literature, which are the scholarly references (journal articles, e-books, and other electronic resources), concerning this study and its interest areas. The PRISMA literature analysis was illustrated in section 2.2, and the literature chart was defined in section 2.3. The electronic word of mouth (eWOM) is depicted in section 2.4, and the influence of eWOM on consumer decisions is demonstrated in section 2.5. In section 2.6, we will look at the previous research regarding how eWOM analysis done by using Natural Language Processing (NLP) and Sentiment Analysis. The importance of product rating is discussed in section 2.7, and separate product rating analyses using eWOM are discussed in section 2.8. Section 2.9 shows the Bidirectional Transformer Encoder Representations (hereafter BERT) Model used in this research. A table of publications based on the literature review can be found in Section 2.10. The systematic literature review's conclusion is summarized in section 2.11.

2.2 PRISMA Literature Review

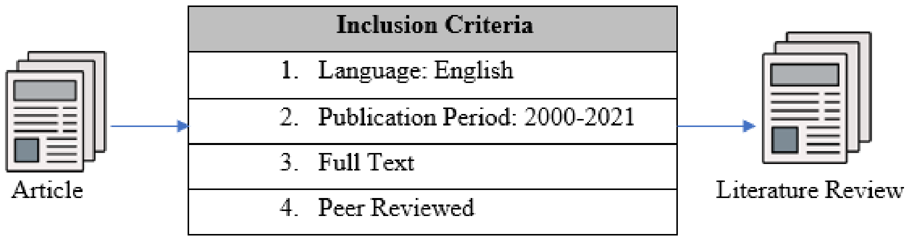
The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standard offers guidelines for Systematic Review reporting. The flow diagram illustrates how material flows through the various stages of a Systematic Analysis. It indicates how many documents were found, which were included, and which were omitted, as well as why they were removed (Moher et al., 2009).

Literature for the literature review is found using the WINTEC digital library's OneSearch tool. OneSearch provides access to several directories, including EBSCO Host, ProQuest, and ScienceDirect. As a result, literature is collected from OneSearch, which provides researchers with free access to many journal papers and other scholarly tools. This study relies on peer-reviewed publications. For the selection and search for the literature, the researcher preferred some keywords like "eWOM analysis", "evaluation of customer reviews", "product rating", and "product ranking based on customer reviews" to get closely related to the research problem. The search for articles also included criteria such as year of publication (between 2000 and 2021), relevance to the subject and language (English).

2.2.1 Inclusion Criteria

When the reviewer receives an article, he considers several inclusion requirements before including it in the literature review. The researcher considers factors such as whether the paper is the full text, whether it is written in English, whether it was published between 2000 and 2021, and whether it was peer-reviewed.

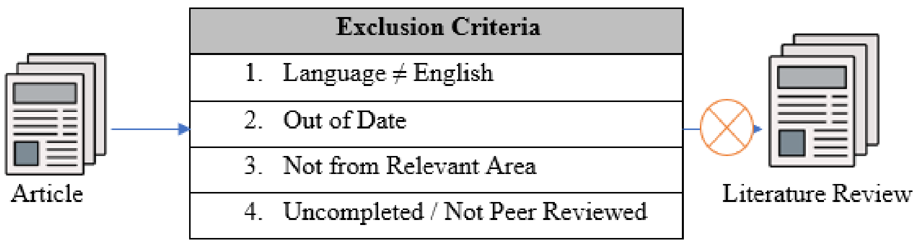
Figure 2.1 - Inclusion Criteria



2.2.2 Exclusion Criteria

To exempt publications from the Literature review, the researcher used parameters such as not accessible, not being written in English, not being from a relevant research area, being out of date based on the appropriate period and being incomplete or not being peer-reviewed.

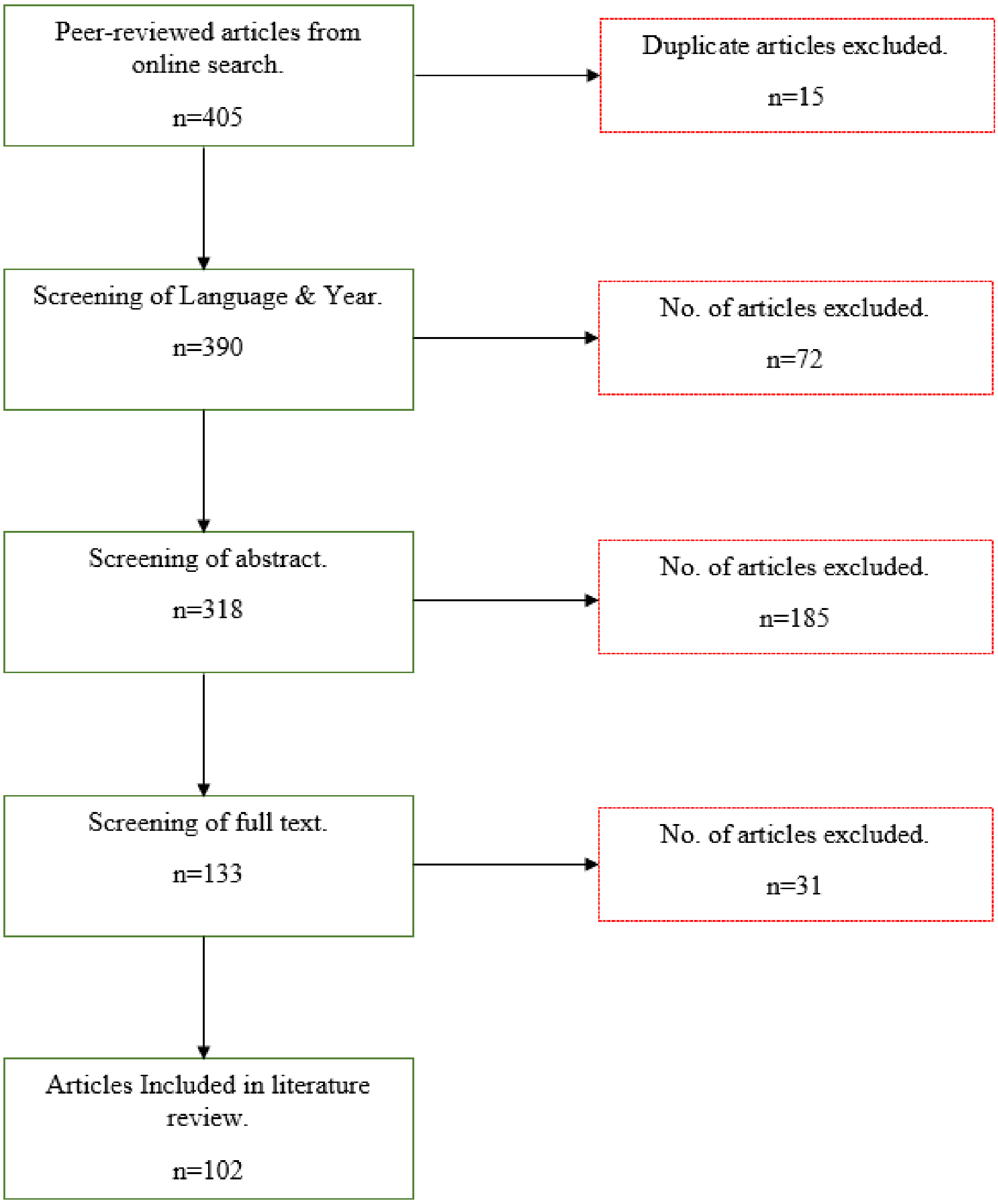
Figure 2.2 - Exclusion Criteria



2.2.3 PRISMA Flow Diagram

This review adheres to the PRISMA guidelines. An analysis of 102 papers written between 2000 and 2021 was conducted. There were 405 papers found using the guided filter. The researcher conducted a preliminary screening of those papers, excluding185 and 15 duplicates, using a description and abstract check. The researcher excludes 72 papers that do not meet the years of publication or the language standards he developed. There were 102 publications left after they were reviewed according to inclusion and exclusion requirements. As a result, there are a total of 102 articles in this study that have been peer-reviewed. Figure 2.3 depicts the PRISMA Flow diagram, which following the inclusion and exclusion criteria.

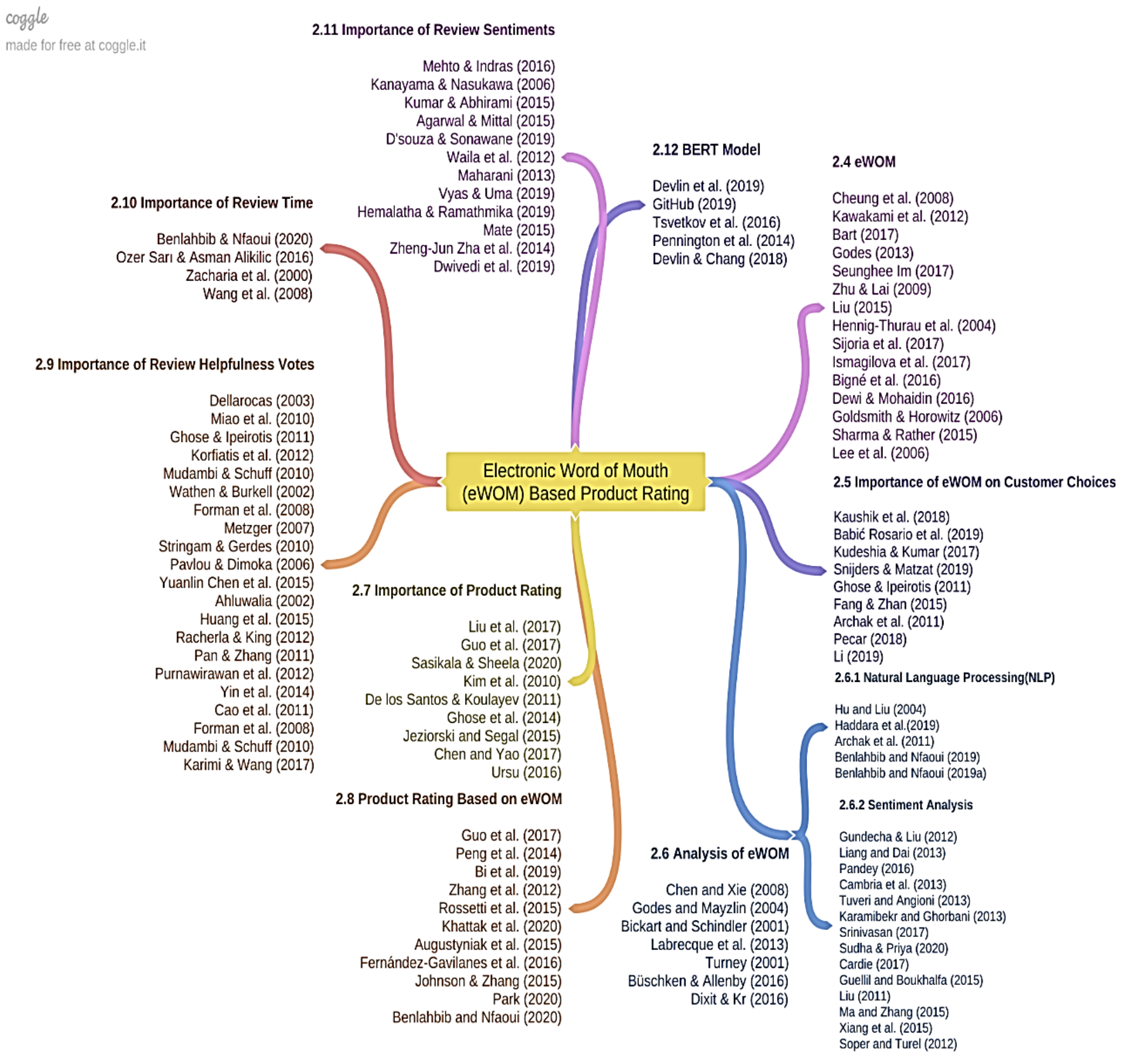
Figure 2.3 - PRISMA Flow Diagram



2.3 Literature Map

To address the shortcomings of conventional literature reviews researcher attempted to cover all references in a comprehensive manner. A concept map represents the literature subject or areas explored in the study and research process (Hlee et al., 2018). The literature map for this study is shown in section 2.3 as Figure 2.4.

Figure 2.4 - Literature Map



2.3 Electronic Word of Mouth (eWOM)

The exchange of sentiments, viewpoints, and opinions between two persons or societies is called Word of Mouth (WoM) and the WoM using online platforms called electronic WoM (Cheung et al., 2008). Word-of-mouth (WOM) has been shown to have a major impact on consumer buying behaviour by influencing customer perception (Kawakami et al., 2012; Bart, 2017; Godes, 2013). In previous studies, Seunghee (2017), Zhu and Lai (2009) and Liu (2015) found that WoM has to be more effective than conventional communication methods such as personal sale and traditional advertising media.

Hennig-Thurau et al. (2004) state that each positive or negative comment made by future, current, or former consumers about a product or business that is made accessible to people and organisations through the Internet is referred to as electronic word-of-mouth (eWOM).

It could be the next step in the development of interpersonal communication into the next age of cyberspace. How eWOM influences consumption has been the subject of numerous marketing and consumer studies (Sijoria et al., 2017; Ismagilova et al., 2017; Bigné et al., 2016). Senecal and Nantel (2004) performed an exploratory study of customers' utilisation of online recommendation channels to see how eWOM influences consumables selection. Other researchers Dewi and Mohaidin (2016), Goldsmith and Horowitz (2006), Sharma and Rather (2015) and Lee et al. (2006) have conducted studies to determine the causes for discovering the effects of eWOM and articulating it, to assist sellers or advertisers in better understanding online customer behaviour. The studies above show that eWOM has become an integral part of the online marketing mix, affecting online consumers' buying decisions significantly.

2.4 Importance of eWOM on Customer Choices.

Online reviews are considered the typical form of a new generation WoM. Online reviews are mainly considered evaluations of a product or service, including some personal feelings and sentiments (Mudambi & Schuff, 2010). Kaushik et al. (2018) identified the significance of online reviews, which have a heavy influence on customer's purchasing decisions than traditional WoM because it never had any geographical barriers to share information regarding different products available in this global market.

Companies are attempting to analyse eWOM to make or alter policies. Marketers are assessing eWOM to ensure customer importance, and as a result, they are trying to deliver customer service to sustain brand equity (Babić Rosario et al., 2019). New modes of communication have emerged as a result of the Internet, and knowledge exchange between customers and businesses, as well as among consumers, has been enabled to the point that the online world has become a privileged forum for exchanging brand experiences (Kudeshia & Kumar, 2017; Snijders & Matzat, 2019). In 2011, Ghose and Ipeirotis found that online reviews by the customers affect the sale of products and play a vital role in purchasing decisions. New customers following these reviews and evaluating the satisfaction level of the product or services with the current user. In addition to this, Fang and Zhan (2015) and Archak et al. (2011) shows evidence of how WoM and online reviews affect customer choices.

Kaushik et al. (2018) performed a holistic approach to the reviews availed from Amazon.com and found that the products that hold more negative reviews can only generate a limited revenue for the company. In addition to this, Pecar (2018) and Li et al. (2019) investigated the impact of feedback using the joint-sentiment topic model and found that online WoM has an influential power over traditional WoM.

2.5 Analysis of eWOM

In 2008, Chen and Xie strongly suggested that customer review evaluation is an important growing technique to evaluate consumer purchase decisions and product sales. Due to the reliability of the information source, consumer-created information in the context of online feedback is deemed more trustworthy than seller-created information. As user-generated product content, online consumer feedback can thus be considered a unique form of word-of-mouth interaction (Godes & Mayzlin, 2004; Bickart & Schindler, 2001). According to researchers, customers gain a great benefit or become inspired shoppers by sharing online feedback (Labrecque et al., 2013).

In 2001, Turney suggested an algorithm for classifying user feedback into two groups (recommended or not). A semanticist orientation between the analysis phrases – positive or negative, based on measured reciprocal point-wise knowledge between the word's pairs, is described there. A subject model for text analysis of consumer feedback is proposed and evaluated on various datasets, considering the numerical ratings provided by users, and it is based on the Latent Dirichlet Allocation (LDA) model (Büschken & Allenby, 2016). In research conducted by Dixit and Santhosh (2016), text classification using Naive Bayes and K-NN classifiers into three groups (excellent, poor, and mixed) was performed using RapidMiner without changing the data model.

2.5.1 Using Natural Language Processing Techniques

Hu and Liu (2004) first introduced a system that summarises customer reviews based on features. However, over the past two decades, so many systems were launched and applied to different domains like a product review, movie review, accommodation review and so on (Haddara et al., 2019; Archak et al., 2011).

In the study of Yan et al. (2017), where comments are combined into separate sets of opinions based on their textual associations, then a standard credit score is created by aggregating the statistics of fused and clustered opinions. Benlahbib and Nfaoui (2019) suggested a fourfold method, and they implemented the K-means clustering algorithm to combine related reviews into the same cluster utilising Latent Semantic Analysis (LSA) before generating a reputational rating for each cluster's statistics.

2.5.2 Using Sentiment Analysis

Sentiment analysis is characterised as the responsibility of finding the suppositions - the sentiments and feelings - of people regarding explicit substances and their extraction (Gundecha & Liu, 2012). Liang and Dai (2013) significantly found that the study of opinions, feelings and sentiment information pulling more attention from both the scientific and business community. The authors such as Pandey (2016), Cambria et al. (2013) and Tuveri and Angioni (2013) not differentiated the specifications of sentiment analysis and text analysis.

On the other hand, Karamibekr and Ghorbani (2013), Srinivasan (2017) and Sudha (2020) contended that conclusion mining and notion investigation could be characterised as an interdisciplinary zone arranged among the fields of Natural Language Processing (NLP). In addition to this, Breck and Cardie (2017) and Liu (2011) distinguished the principles between text mining and sentiment analysis.

Guellil and Boukhalfa, 2015) describe the distinctive nature or features of opinion mining as three groups: modelling, extraction, and subjectivity analysis of opinions. Modelling deals with how the opinion becomes formalised; extraction distinguishes the subjects such as general, multiple subjects, and the person who registered the opinions. If the review or viewpoint or opinion contains any facts, it will be categorised as objective, and the opinion holds any personal information that is categorised under subjective (Liu, 2011).

Ma and Zhang (2015) studied the relationship between predictive analytics and the feelings of the public. They extracted the data, which includes the public's feelings, from well-reputed social media and e-business companies named Sina Microblog and Taobao. They used F-statistics (significance), MAPE (Means Absolute Percentage Error and R-square) and concluded that the company's revenue increased to the upbeat mood of the public.

Apart from other studies, Xiang et al. (2015) made an exception that they perform a text mining approach to evaluate the reviews in Expedia.com. They used frequency measuring of each word using statistical methods to gather the comfort level of the customer. The authors take single words than the combination of words and sentences for their evaluation process (Xiang et al., 2015). So, the n-gram analysis suggested by Soper and Turel (2012) did not follow by Xiang et al. (2015).

2.6 Importance of Product Rating

There haven't been many reports on how to rate products based on online feedback. There is only a small amount of literature that explicitly or indirectly discusses this issue. The studies conducted by Liu et al. (2017), Guo et al. (2017) and Sasikala and Sheela (2020) have made significant contributions to rating products through online reviews. The causal effect of rating and search costs on customer preference is empirically studied in a series of articles in the marketing literature (Kim et al. 2010, De Los Santos & Koulayev, 2011; Ghose et al., 2014; Jeziorski & Segal, 2015; Chen & Yao, 2017; Ursu,2016). All the experiments showed that putting an item in high-visibility areas raises the chance of it being bought. Another result that seems to be common in this line of research is that w-ordered ranking increases consumer protection as opposed to commonly used ranking approaches. These studies do not argue, however, that the w-ordered rating maximises consumer welfare.

2.7 Product Rating through online product reviews

Many studies use text mining (aspect-based sentiment information) techniques from online feedback (Guo et al., 2017; Bi et al., 2019). Zhang et al. (2012) manually defined a collection of synonyms, based on which they identified the sentiment polarity of the entire analysis and modelled by creating a guided and weighted graph, on which they were able to rate items.

Rossetti et al. (2015) followed a text mining way to deal with recommendations to the customers based on available online reviews. They implemented topic modelling based on the frequency of words to make recommendations. The Latent Dirichlet allocation (LDA) used to analyse the distribution of words regarding different topics in the customer reviews, and the authors claimed it is one of the most accepted modelling methods across the globe (Rossetti et al., 2015). Pang et al. (2002) proposed a regulated report level sentiment mining utilising a learning classifier with various surveys about films. Khattak et al. (2020) used Fuzzy-based sentiment analysis to classify the customer sentiments at a fine-grained level from their reviews by extending fuzzy hedges along with its rule-sets. Finally, they crossed the success line to demonstrate improved performance and results using an extended set of Fuzzy linguistic hedges. Augustyniak et al. (2015) compared different classifiers like Random Forests, Linear SVC, Bernoulli Naïve Bayes, Multinomial Naïve Bayes, Extra Tree Classifier, Logistic Regression and AdaBoost and suggest that Logistic Regression outperforms in sentiment reviews polarity prediction. Fernández-Gavilanes et al. (2016)

found that online textual messages depend more on unsupervised dependency parsing-based text classification method. Johnson and Zhang (2015) presented an additional Convolutional Neural Network (CNN) in deep learning methods and measured that 92.33% of accuracy.

Park (2020) used the dictionary-based approach using WordNet and the corpus-based approach. Consequently, an emotion score is given to each sentence, and the viewpoint of each sentence is calculated as optimistic, negative or neutral. Benlahbib and Nfaoui (2020) followed the BERT model to define sentiment and relations between words in a sentence, and this fine-tuned BERT provided the probability of being negative or positive, and they apply a max function to the output of BERT.

2.8 Importance of Review Helpfulness

Consumers who have already bought such goods create reviews that explain consumers' perceptions of products, services or related knowledge in general practice. The widespread availability of product reviews mainly on the world wide web has turned them into one of the most valuable sources of knowledge for consumers making purchasing decisions (Dellarocas, 2003; Miao et al., 2010). According to various research, consumer votes on the helpfulness of written reviews, making a critical impact in affecting purchasing decisions (Ghose & Ipeirotis, 2011; Mudambi & Schuff, 2010; Korfiatis et al., 2012). Customer reviews with high favourable or unfavourable grades are bagged as more insightful and, therefore, beneficial. In addition, long and robust reviews can encourage confidence and are considered beneficial (Wathen & Burkell, 2002; Forman et al., 2008; Metzger, 2007; Stringam & Gerdes, 2010). Given the theory of the information diagnosticity theory, a user review's helpfulness factor depends on whether a

review could minimise its confusion when making buying decisions (Feldman & Lynch, 1988 as cited in Yuanlin Chen et al., 2015). The study conducted by Yuanlin Chen et al. (2015) defined review helpfulness as the ratio in between the total number of users who find the review is beneficial and marked as 'helpful' and the total number of users who read this user review. Numerous review and reviewer attributes that affect review helpfulness have been identified in previous studies, and user review length, rating valence, and the review extremity are the common review attributes for review helpfulness (Huang et al., 2015; Racherla & King, 2012; Pan & Zhang, 2011; Purnawirawan et al., 2012; Yin et al., 2014). User review length, rating valence, and the review extremity are the common review attributes for review helpfulness (Cao et al., 2011; Forman et al., 2008; Mudambi & Schuff, 2010). In addition, Karimi & Wang (2017) studied the reviewers' innovation, identity disclosure, reviewers' skills and credibility.

The source's authority has shown a significant impact on the public authenticity of the customer reviews. Likewise, the studies suggest that the credibility of their contributor may partly influence the reactions of consumers to a given review. As a result, reviews written by prominent reviewers may directly affect product or service revenue (Wathen & Burkell, 2002; Pavlou & Dimoka, 2006; Forman et al., 2008; Ghose & Ipeirotis, 2011). So, in this research, the researcher considering the evaluation and scoring of the review helpfulness as one of the influencing factors which will improve the accuracy in an automated review and product rating model.

2.9 Importance of Review Time

Benlahbib and Nfaoui (2020) introduced a study for the visual reputation of online reviews, which considered the review time as one of the most influenced factors in customer review evaluation. According to Sari and Allkilic (2016), technology causes consumers to lose patience to some extent. If a consumer is disappointed, they will not hesitate to post a scathing review of the company's product or service on online forums in an immediate effect. The studies by Zacharia et al. (2000) and Wang et al. (2008) show that review ageing is a serious problem that must be properly addressed when used in the review analysis process. Only a few studies have looked at review time as a consideration in a product evaluation or review interpretation. However, in this analysis and the proposed product rating model, the researcher took review time into account as one of the main factors which will improve the accuracy of an automated review and product rating model.

2.10 Importance of Review Sentiments

A sentiment can be described as a user's offensive comment. Sentiments are any user's opinions, such as liking or desiring something (positive), disliking or undesiring something (negative), and sometimes it may be a neutral one (Mate, 2015; Zheng-Jun Zha et al., 2014; Dwivedi et al., 2019). Sentiment analysis of consumer reviews will reveal the user's feelings about a product. The majority of the studies focus on finding the aspects of a review that direct its polarity. Machine learning and semantic orientation methods are two broad categories of techniques used for sentiment analysis models (Kanayama & Nasukawa, 2006; Kumar & Abhirami, 2015; Agarwal & Mittal, 2015; D'souza & Sonawane, 2019;

Waila et al., 2012; Maharani, 2013; Vyas & Uma, 2019; Hemalatha & Ramathmika, 2019). A study conducted by Mehto & Indras (2016) on lexicon-based emotion analysis demonstrates the importance of sentiment analysis for text analysis activities. So, in this research, the researcher considering the evaluation and scoring of the review sentiment as one of the influencing factors which will improve the accuracy of an automated review and product rating model.

2.11 BERT Model

BERT is a Natural Language Processing Model introduced by researchers at Google Research in 2018. BERT depends on Transformers, a profound learning model in which each input component is associated with each output component, and the weightings between them are progressively determined dependent on their association. It is a modern language representation paradigm built to pre-train deep bi-directional representations of unmarked text by jointly conditioning both the left and the right in both layers (Devlin et al., 2019). To find the proper meaning of a language, the bidirectionality of a model is essential. Figure 2.5 and Figure 2.6 illustrates the BERT language model and Elmo model, respectively.

BERT regarded as the first unsupervised, deeply bidirectional method for pre-training NLP. Therefore, it works better than the previous models. Unsupervised implies that BERT's training process used just a plain content corpus, which is significant because lots of plain text are freely accessible on the web in numerous languages (GitHub, 2019). A term from the embedding layer begins with its embedding representation. To develop a new intermediate representation, each layer just does a multi-headed focus calculation on the word representation of the previous layer, and these intermediate layers are in the same size. In Figure 2.5, E is the embedding representation, representing the same tokens marked with Tom as T and intermediate parts. If it is a BERT model of 12 layers, so it should have 12 intermediate layers (Devlin et al., 2019).

BERT uses the word Piece tokenisation. For all the actual characters in the language, the vocabulary is initialised and then iteratively incorporates the most frequent/likely variants of the current words in the vocabulary. All the pre-training objectives of BERT allow the usage of BERT in a wide variety of tasks (GitHub, 2019).

Figure 2.5 - BERT Architecture (Devlin et al., 2019)

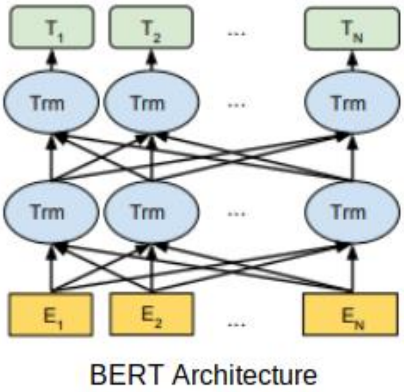
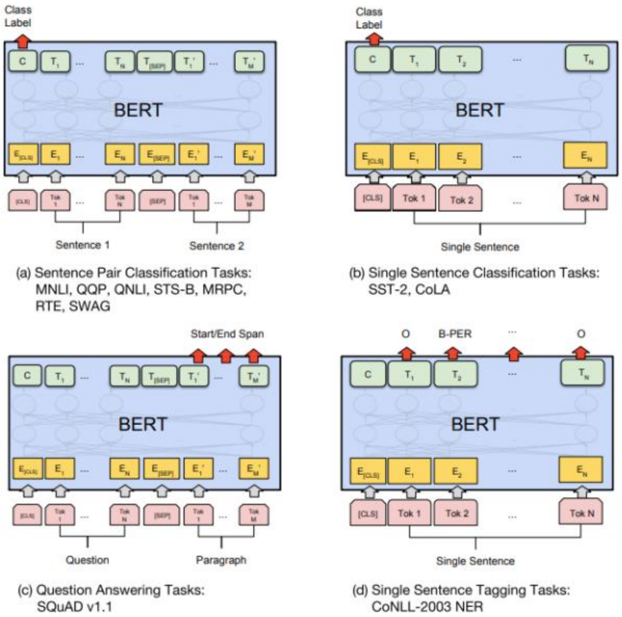


Figure 2.6 - BERT language model (Devlin et al., 2019)



The fine-tuning procedure of the BERT model for sequence classification tasks has a fixed representation denoted as [CLS], considered as the hidden state of the token added by the classification layer. The addition of parameter follows a dimension K x H, where K means the number of classifier labels, H stands for the size of the hidden state and finally, the standard SoftMax method is used to compute the probabilities of labels. The fine-tuning procedure for the sentence pair classification tasks method is directly comparable to the role of single sequence classification. In the input representation, where the two sentences are concatenated together, the only difference is.

2.12 What distinguishes BERT?

The representations, especially the pre-trained ones, can be context-free or contextual, and these representations may be unidirectional or bi-directional also. For each word in the language, context-free models like word2vec (a neural network that handles text data, as introduced by Google) or GloVe (a method for obtaining vector representations for terms using unsupervised learning algorithm)(Pennington et al., 2014) create a word embedding representation. In context-free representation, the word "bank" may be used for "bank account" or "bank of the river". Instead, contextual models describe each word depending on the other words in the sentence. However, from the very ground of a deep neural network, BERT represents a "bank," making it profoundly bidirectional, using both its previous and its next meaning (Devlin & Chang, 2018). BERT is the first time a deep neural network has been successfully pre-trained (Devlin et al., 2019).

In 2018, Devlin & Chang contrasted BERT to other advanced NLP systems to test its performance. Importantly, with almost no specific improvements to the neural network design, BERT has accomplished all its outcomes. Compared with some other NLP models, the BERT scoring is 93.2% F1, which is above the current state-of-the-art scoring of 91.6% and human scoring of 91.2%.

2.13 List of Articles in Literature Review

Table 2.1 shows a list of literature related to the topic discussed in this Literature Review chapter.

Table 2.1 - List of articles in Literature Review

Topic	Article	Author(s)
WOM & eWOM Definition	"The impact of electronic word-of-mouth"	Cheung et al. (2008)
	"Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet?"	Thurau et al. (2004)
Impact of WOM on customer purchase preferences	"Product seeding: Word-of-Mouth effects for and beyond the focal product"	Bart (2017)

	"Product policy in markets with word-of-Mouth communication"	Godes (2013)
	"Personal word of mouth, virtual word of mouth, and innovation use"	Kawakami et al. (2012)
	"Online WOM effects and product type: Evidence from Tmall"	Liu (2015)
	"The relationship between consumers' WOM motivations and the valence of WOM on movie"	Seunghee Im (2017)
	"A study about the WOM influence on tourism destination choice"	Zhu and Lai (2009)
How eWOM effect the selection of consumables	"The influence of online product recommendations on consumers' online choices"	Senecal and Nantel (2004)
	"Impact of eWOM"	Ismagilova et al. (2017)
	"What makes eWOM viral?"	Sijoria et al. (2017)
	"EWOM on travel agency selection: Specialized versus private label"	Bigné et al. (2016)
To determine the effects of eWOM and to assist sellers or advertisers in better understanding online customer behaviour.	"Motivations of online opinion seeking and its effect on the online purchase intention"	Dewi and Mohaidin (2016)
	"Measuring motivations for online opinion seeking	Goldsmith and Horowitz (2006)
	Understanding customer knowledge sharing in web-based discussion boards"	Lee et al. (2006)

	"Understanding the customer experience: An exploratory study of a category hotels"	Sharma and Rather (2015)
eWOM influence	"Research note: What makes a helpful online review? A study of customer reviews on Amazon.com"	Mudambi and Schuff (2010)
	"Exploring reviews and review sequences on e-Commerce platform: A study of helpful reviews on Amazon.in"	Kaushik et al. (2018)
Aim of eWOM analysis	"Conceptualizing the electronic word-of-mouth process: What we know and need to know about eWOM creation, exposure, and evaluation"	Babić Rosario et al. (2019)
	"Social eWOM: Does it affect the brand attitude and purchase intention of brands?"	Kudeshia and Kumar (2017)
	"Online reputation systems"	Snijders and Matzat (2019)
	"Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics"	Ghose and Ipeirotis (2011)
Impact of eWOM	"Sentiment analysis using product review data"	Fang and Zhan (2015)
	"Deriving the pricing power of product features by mining consumer reviews"	Archak et al. (2011)

	"Exploring reviews and review sequences on e-Commerce platform: A study of helpful reviews on Amazon.in"	Kaushik et al. (2018)
	"Towards opinion summarization of customer reviews"	Pecar (2018)
	"The effect of online reviews on product sales: A joint sentiment-topic analysis"	Li et al. (2019)
eWOM Analysis	"Online consumer review: Word-of-Mouth as a new element of marketing communication mix"	Chen and Xie (2008)
	"Using online conversations to study word-of-Mouth communication"	Bickart and Schindler (2001)
	"Internet forums as influential sources of consumer information"	Bickart and Schindler (2001)
	"Consumer Power: Evolution in the Digital Age"	Labrecque et al. (2013)
	"Sentence-based text analysis for customer reviews"	Büschken and Allenby (2016)
	"Thumbs up or thumbs down? Thumbs up or thumbs down?"	Turney (2001)
	"Collaborative analysis of customer feedbacks using rapid miner"	Dixit and Santhosh (2016)

Natural Language Processing (NLP) Techniques	"Exploring customer online reviews for new product development: The case of identifying reinforcers in the cosmetic industry"	Haddara et al. (2019)
	"Deriving the pricing power of product features by mining consumer reviews"	Archak et al. (2011)
	"Prediction (or not) during language processing"	Yan et al. (2017)
	"A hybrid approach for generating reputation based on opinions fusion and sentiment analysis"	Benlahbib and Nfaoui (2019)
	"An unsupervised approach for reputation generation"	Benlahbib and Nfaoui (2019a)
Sentiment Analysis	"Mining social media: A brief introduction"	Gundecha and Liu (2012)
	"Opinion mining on social media data"	Liang and Dai (2013)
	"Sentiment analysis through text Mining-A review"	Pandey (2016)
	"New avenues in opinion mining and sentiment analysis"	Cambria et al. (2013)
	"An opinion mining model for generic domains"	Tuveri and Angioni (2013)
Sentiment Analysis and Natural Language Processing	"A structure for opinion in social domains"	Karamibekr and Ghorbani (2013)
	"Marketing applications using big data"	Srinivasan (2017)
	"Social media sentiment analysis for opinion mining"	Sudha (2020)

Difference between Text Mining and Sentiment Analysis	“Opinion mining and sentiment analysis”	Breck and Cardie (2017)
	“Opinion mining and sentiment analysis”	Liu (2011)
	“Social big data mining: A survey focused on opinion mining and sentiments analysis”	Guellil and Boukhalfa (2015)
Methods used for Sentiment Analysis	“Public mood and consumption choices: Evidence from sales of Sony cameras on Taobao”	Ma and Zhang (2015)
	“What can big data and text analytics tell us about hotel guest experience and satisfaction?”	Xiang et al. (2015)
	“Who are we? Mining institutional identities using N-grams”	Soper and Turel (2012)
	“Social big data mining: A survey focused on opinion mining and sentiments analysis”	Guellil and Boukhalfa (2015)
Importance of Product Ranking	“Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory”	Liu et al., (2017)
	“Mining online customer reviews for products aspect-based ranking”	Guo et al. (2017)
	“Sentiment analysis of online product reviews using	Sasikala and Sheela (2020)

	DLMNN and future prediction of online product using IANFIS”	
	“Online demand under limited consumer search”	Kim et al. (2010)
	“Optimizing click-through in online rankings for partially Anonymous consumers”	De los Santos and Koulayev (2011)
	“Examining the impact of ranking on consumer behavior and search engine revenue”	Ghose et al. (2014)
	“What makes them click: Empirical analysis of consumer demand for search advertising”	Jeziorski and Segal (2015)
	“Sequential search with refinement: Model and application with click-stream data”	Chen and Yao (2017)
	“The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions”	Ursu (2016)
Product Ranking from eWOM	“Mining online customer reviews for products aspect-based ranking”	Guo et al. (2017)
	“Representing sentiment analysis results of online reviews using interval type-2 fuzzy numbers and its	Bi et al. (2019)

	application to product ranking”	
	“Mining millions of reviews”	Zhang et al. (2012)
	“Analysing user reviews in tourism with topic models”	Rossetti et al. (2015)
	“Fine-grained sentiment analysis for measuring customer satisfaction using an extended set of fuzzy linguistic hedges”	Khattak et al. (2020)
	“Comprehensive study on lexicon-based ensemble classification sentiment analysis”	Augustyniak et al. (2015)
	“Unsupervised method for sentiment analysis in online texts”	Fernández-Gavilanes et al. (2016)
	“Effective use of word order for text categorization with Convolutional neural networks”	Johnson and Zhang (2015)
	“Framework for sentiment-driven evaluation of customer satisfaction with cosmetics brands”	Park (2020)
	“Aggregating customer review attributes for online reputation generation”	Benlahbib and Nfaoui (2020)
Importance of Review Helpfulness	“The digitization of word of mouth: Promise and challenges of online feedback mechanisms”	Dellarocas (2003)

	"Journal of the American Society for Information Science and Technology"	Miao et al. (2010)
	"Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics"	Ghose and Ipeirotis (2011)
	"Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content"	Korfiatis et al. (2012)
	"Research note: What makes a helpful online review? A study of customer reviews on Amazon.com"	Mudambi and Schuff (2010)
	"Believe it or not: Factors influencing credibility on the web"	Wathen and Burkell (2002)
	"Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets"	Forman et al. (2008)
	"Making sense of credibility on the web: Models for evaluating online information and recommendations for future research"	Metzger (2007)
	"An analysis of word-of-Mouse ratings and guest	Stringam and Gerdes (2010)

	comments of online hotel distribution sites”	
	“The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation”	Pavlou and Dimoka (2006)
	“Analysis of review helpfulness based on consumer perspective”	Yuanlin Chen et al. (2015)
	“How prevalent is the negativity effect in consumer environments?”	Ahluwalia (2002)
	“A study of factors that contribute to online review helpfulness”	Huang et al. (2015)
	“What we know and don’t know about online word-of-Mouth: A systematic review and synthesis of the literature”	Racherla and King (2012)
	“Born unequal: A study of the helpfulness of user-generated product reviews”	Pan and Zhang (2011)
	“Balance and sequence in online reviews: How perceived usefulness affects attitudes and intentions”	Purnawirawan et al. (2012)
	“Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews”	Yin et al. (2014)

	"Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach"	Cao et al. (2011)
	"Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets"	Forman et al. (2008)
	"Research note: What makes a helpful online review? A study of customer reviews on Amazon.com"	Mudambi and Schuff (2010)
	"Online review helpfulness: Impact of reviewer profile image"	Karimi and Wang (2017)
Importance of Review Time	"Aggregating customer review attributes for online reputation generation"	Benlahbib and Nfaoui (2020)
	"How ready are the Turkish hospitality and travel organizations for E-complaint handling?"	Ozer Sarı and Asman Alikilic (2016)
	"Collaborative reputation mechanisms for electronic marketplaces"	Zacharia et al. (2000)
	"Improving the Amazon review system by exploiting the credibility and time-decay of public reviews"	Wang et al. (2008)
Importance of Review Sentiment	"Data mining through sentiment analysis: Lexicon based sentiment analysis	Mehto and Indras (2016)

	model using aspect catalogue”	
	“Fully automatic lexicon expansion for domain-oriented sentiment analysis”	Kanayama and Nasukawa (2006)
	“An experimental study of feature extraction techniques in opinion mining”	Kumar and Abhirami (2015)
	“Semantic orientation-based approach for sentiment analysis”	Agarwal and Mittal (2015)
	“Sentiment analysis based on multiple reviews by using machine learning approaches”	D'souza and Sonawane (2019)
	“Evaluating machine learning and unsupervised semantic orientation approaches for sentiment analysis of textual reviews”	Waila et al. (2012)
	“Microblogging sentiment analysis with lexical based and machine learning approaches”	Maharani (2013)
	“Approaches to sentiment analysis on product reviews”	Vyas & Uma (2019)
	“Sentiment analysis of yelp reviews by machine learning”	Hemalatha and Ramathmika (2019)
	“Survey on product review sentiment analysis with aspect ranking”	Mate (2015)

	“Product aspect ranking and its applications”	Zheng-Jun Zha et al. (2014)
	“Movie reviews classification using sentiment analysis”	Dwivedi et al. (2019)
Bidirectional Transformer Encoder Representations (BERT) Model	“BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”	Devlin et al. (2019)
	“Correlation-based intrinsic evaluation of word vector representations”	Tsvetkov et al. (2016)
	“Glove: Global vectors for word representation”	Pennington et al. (2014)
	“Open sourcing BERT: State-of-the-Art pre-training for natural language processing”	Devlin and Chang (2018)

2.14 Research Gap

The researcher followed previous studies and articles on product rating based on consumer review data to determine the study issue. Then, based on the written literature, especially discussed in section 2.6, section 2.7 and section 2.8, the researcher determined that there isn't a solution that considers things like review sentiment, review validity, and reviews helpfulness. The researcher looked into the topic's validity and concluded that the analysis regarding a new solution for product ranking is relevant when the considerations mentioned above are taken into account.

2.15 Main Research Question

According to Tuckman and Harper (2012), Ratan et al. (2019) and Creswell (2014), A research question expresses your interest in a certain subject or phenomenon. Whether the research is descriptive or experimental, identifying a research question will help to focus the research or define the path of the investigation. If the study is experimental, defining a research question may help to focus the research or explain the course of the enquiry (Tuckman & Harper, 2012; Ratan et al., 2019). As a result of this literature review, which leads to a research gap mentioned in section 2.15, the researcher formulated the main research question.

> **RQ 1**: What is the importance of attributes associated with online customer reviews to define the accuracy of product rating?

In addition to this, the sub-questions and hypothesis for this study will be explained in Section 3.5.

2.16 Conclusion

Since all previous research primarily based on either the text or the sentiment of consumer feedback, the PRISMA literature review on 72 previous publications found an opportunity to build a product rating system by properly analysing online reviews. The research methodology, research questions and hypothesis, and research design are all discussed in Chapter 3.

3. METHODOLOGY

3.1 Introduction

The research design is presented in this chapter. The topic depicts the overall research design and the research approach, which is accompanied by the study. Pruzan (2016) defines research methodology as the analysis of methods used to find answers to a research issue. The methodology of research is combined with the assessment and examination of utilized strategies for research. To explain the subject, find answers to research questions, and decide the relationship between e-commerce company and customer feedback, science and communication theories were used in this analysis. WOM has evolved, and technology has the relationship between consumer and company, posing new obstacles for service providers to list products according to absolute ratings.

Section 3.2 explains the objective of the research, and section 3.3 depicts the philosophical worldview using in this research. Section 3.4 introduces the design of this research, which provides a summary of the theoretical structure and how the variables are arranged and connected. The research questions and hypothesis are outlined in Section 3.5. The research model, research approach, data collection and analysis was addressed in Section 3.6. Sections 3.7, 3.8, and 3.9, respectively, show the data processing process, main data definition, and data interpretation method. Finally, the conclusion is presented in Section 3.10 of this chapter.

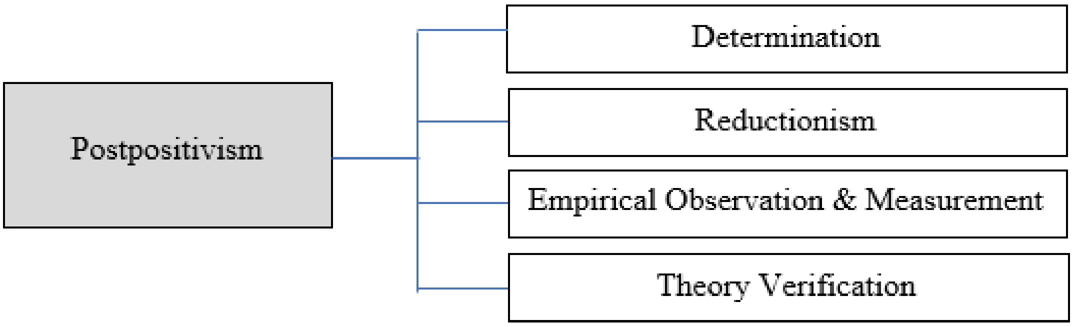
3.2 Objective of the Research

Once again, the researcher stating the objective of this particular research in this section. More of the previous studies mainly focused on polarity detection categorised as supervised and lexicon-based methods on customer reviews (Kanayama and Nasukawa, 2006; Kanayama et al., 2011). Saldaña (2018) identified and proved that a combination of general conceptual semantics with supervised classifiers improved the efficiency of sentiment analysis. Studies conducted by Breck and Cardie (2017) supports opinion mining, and one of the other researchers, Zehra et al. (2017), proposed an ontology-based sentiment analysis. These studies are primarily focused on using semantic and sentiment analysis, disregarding other useful information that could be extracted from user reviews, such as "reviewed date", plays a vital role to analyse the up-to-date information and "helpful reviews", which helps to identify most extreme conditions of the product or service along with sentiment analysis of customer reviews. For that reason, the researcher proposes a sentiment analysis-based rating that incorporates all these attributes during the process of generating rating for various products.

3.3 Philosophical Worldview

Creswell (2014) used the term worldview as meaning "a basic set of beliefs that guide action" (Guba, 1990, p. 17 cited in Creswell (2014)). For research, there have been four philosophical worldviews. However, this study's theoretical framework was post-positivism. Postpositivist believes in determinism, which may be a cause to decide effects or results. Therefore, the researcher chose a post-positivism worldview for this study. Figure 3.1 outlines the Postpositivism worldview.

Figure 3.1 - Postpositivism Philosophy Worldview



3.4 Research Methodology

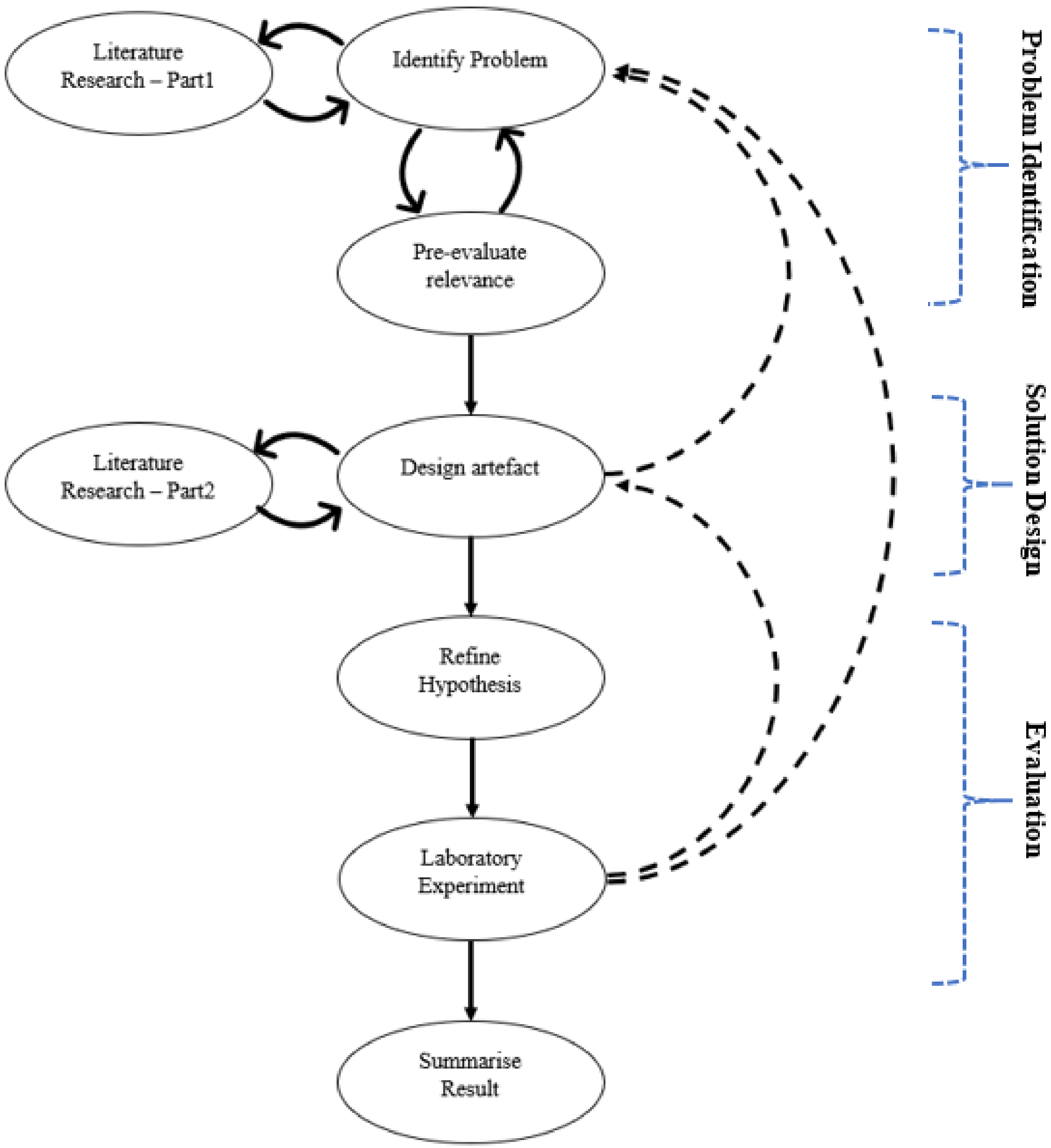
According to Creswell (2013), the research approach is the process and preparation for the study that details the research methodology and performs interpretation and data collection. Overall decisions, philosophical forecasts, a data processing analysis process, and inquiries are all part of a research strategy. The method chosen should be determined by the study topic being discussed. Quantitative analysis, qualitative research, and mixed research are the three types of methods. The quantitative approach looks at how the factors are related. This investigation is focused on an experiment. As a result, a quantitative experimental approach is the right choice for this study.

Researchers use research methodology to learn about how to investigate a particular concept or interest. It may also be described as an action plan, policy, mechanism, or design that directs the collection of methods and links them to the research objectives (Alturki et al., 2011). Since it is an approach that deals with individual, organisational, and social problem solving through artefact creation, the Design Science Research (from now on referred to as DSR) Methodology has seen much success as a method in Information Science and Computer Science (Hevner & Chatterjee, 2010). DSR is a comparatively recent research approach to solve challenges rather than describing or making sense of a fact (Reubens, 2019; Aken, 2004).

According to Offermann et al. (2009), the DSR is divided into three major phases: (a) problem identification, (b) solution design, (c) evaluation and all of which engage with one another during the testing process. Apart from this, each phase is divided into different steps also. Figure 9 depicts the DSR research process. In identify problem step, a problem must be identified, and the literature review process is used to do so. Once a satisfactory issue has been found, a pre-evaluation of significance must be carried out, which is currently being done on the pre-evaluate relevance level. The solution is designed in the second process. It is split down into two parts: (a) design of artefact (b) literature review. Artefact design is a form of innovative engineering. The issue could be re-stated during the artefact design phase. The emphasis of this literature review should be on scientific articles that are important (Hevner & Chatterjee, 2010). The assessment of the solution will begin until it has reached a satisfactory condition. The assessment of this study would be done by following experiments or simulations.

The researcher followed previous studies and articles on product rating based on consumer review data to determine the study issue. Then, based on the written literature, he determined that no solution considers things like review sentiment, review validity, and review helpfulness. The researcher looked into the topic's validity and concluded that the analysis regarding a new solution for product ranking is relevant when the considerations mentioned above are taken into account. From different studies (Wieringa, 2014; Reubens, 2019; Aken, 2004; Offermann et al., 2009), the researcher found that the DSR model is one of the most appropriate models for this research. This study following the six steps designed by DSR, such as identifying the problem, pre-evaluate the relevance of that particular problem, followed by the artefact's design. To refine the hypothesis of this research is mandatory to finetune the research result. Then, the laboratory experiment section will examine the solution, and finally, this study will sum up the results. The researcher strongly believes that this DSR model's iterative capacity will help him restart the process if it fails in a laboratory experiment or the design artefact. Both literature analysis and a pre-evaluation of significance would aid the researcher in fully identifying the issue before beginning their research. The study will report its findings in the final stage of this research model to communicate what happened during the study (Hevner & Chatterjee, 2010; Baskerville et al., 2015). The DSR model supports surveys, laboratory experiments, and case studies but here, this study following the laboratory experiment method, so the researcher modified the DSR research model as shown in Figure 3.2 for this study.

Figure 3.2 - DSR (modified by the researcher from Offermann et al., (2009))



3.5 Research Questions and Hypothesis

This research deals with the following main research question and dealing with three sub-questions that focus on analysing the result of this study.

RQ 1: What is the importance of attributes associated with online customer reviews to define the accuracy of product rating?

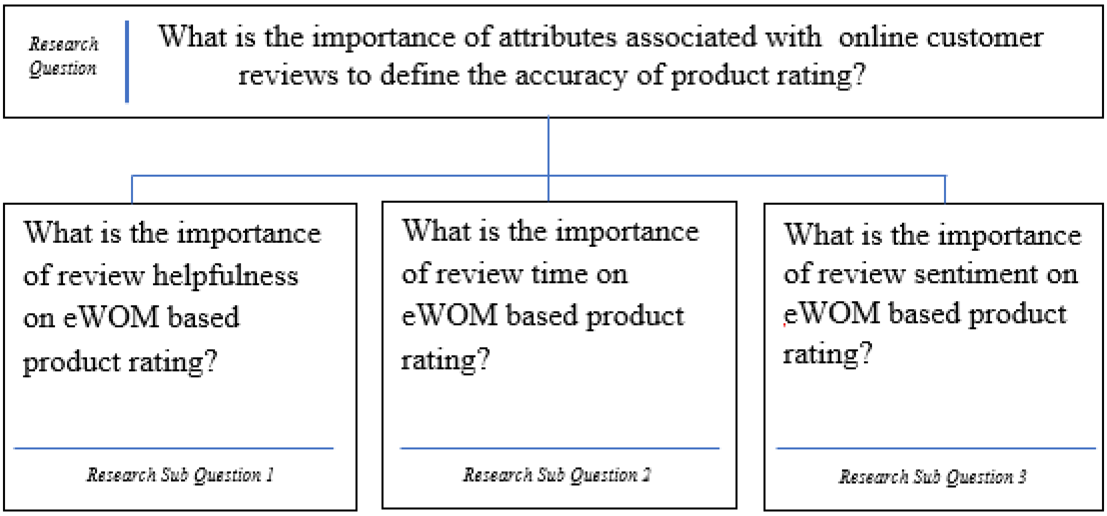
RQ 1.1: What is the importance of review helpfulness on eWOM based product rating?

RQ 1.2: What is the importance of review time on eWOM based product rating?

RQ 1.3: What is the importance of review sentiment on eWOM based product rating?

Figure 3.3 depicts the research question and sub-questions used in this study.

Figure 3.3 - Research Question and Sub Questions for this Study



To answer the research questions, this research aimed at a model, an explanatory model, to explain the association between the product rating and review sentiments and the helpfulness and reviewed time. Regression analysis supports a straightforward explanation of how the dependent variable reacts to the change in the independent variables (Myers, 1990), but the application of regression analysis raises some methodological issues associated with the overall rating of the products. In real-world experiences, consumers participate in an appraisal process for a suitable response, so appraisal theory plays a vital role in all the researches focused on customer reviews and their sentiments. Based on appraisal theory, the previously written reviews will affect the next reviewer's sentiments and their own experience. To find a better method, this research will focus on three variables: review helpfulness, review sentiment, and time of the review by the reviewer. Hence, as discussed in Chapter 2, the hypotheses for this research are as follows:

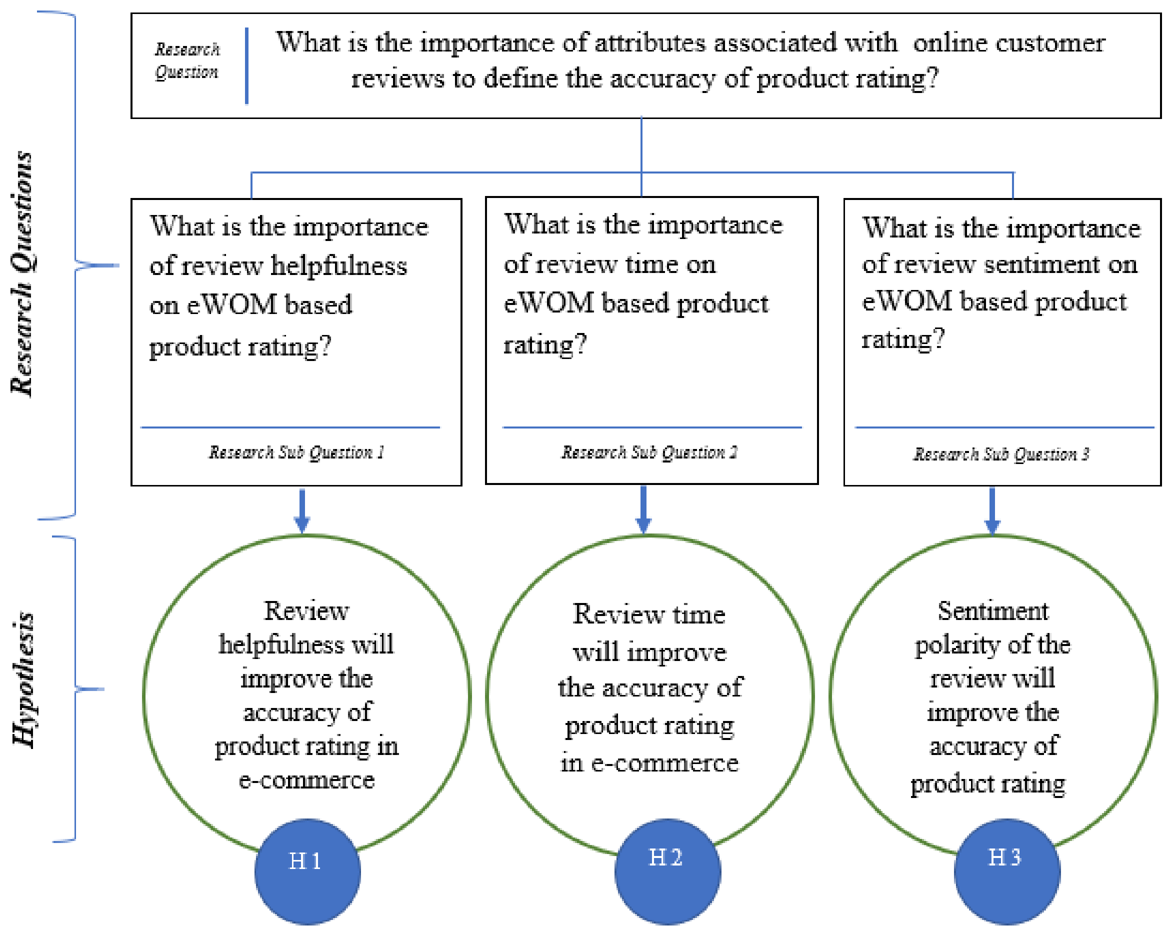
Hypothesis 1 (H1): The accuracy of product ranking in e-commerce would increase if the helpfulness votes of reviews are considered.

Hypothesis 2 (H2): The accuracy of product ranking in e-commerce would increase if the helpfulness votes of reviews are considered.

Hypothesis 3 (H3): The accuracy of product ranking in e-commerce would increase if the helpfulness votes of reviews are taken into account.

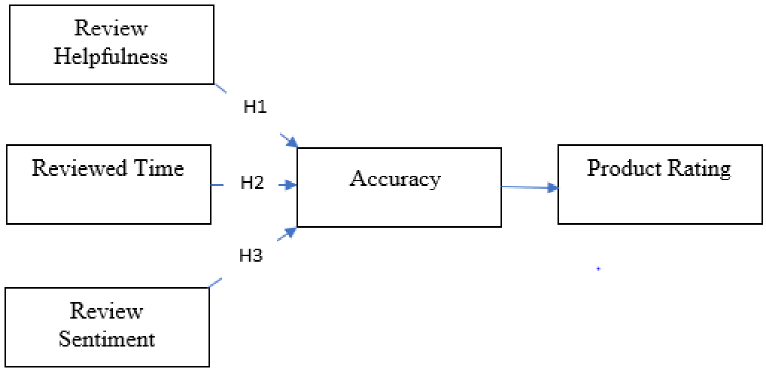
Figure 3.4 depicts the study's hypotheses as well as its research questions.

Figure 3.4 - Research Questions and Hypotheses



In this study, the researcher using different dependent and independent variables to prove the relationship between them. Figure 3.5 depicts the relationship between variables and hypothesis in this study.

Figure 3.5 - Connection of variables and hypothesis



In Figure 3.5, there are three independent variables (Review Helpfulness, Reviewed Time and Review Sentiment) tested using three hypotheses. Review helpfulness means the corresponding value to the number of users who marked it as a helpful review. In the case of reviewed time, this study considers the relevance of reviewed time to calculate a relevant product rating. For instance, if the product is launched in 2010 when the user marked a positive or negative comment in 2020, it has only less weightage than the years closer to the product launch because the specification for the product is relevant to the year 2010. Finally, the review sentiment variable means the result of BERT sentiment analysis. The following tables, Table 3.1 and Table 3.2 describing the relationship between the main research question and literature review and sub research questions, hypothesis and literature review, respectively.

Table 3.1 - Relation between the Main Research Question and Literature Review

Main Research Question	**Literature Review**
RQ1	2.6 2.7 2.8 2.9 2.10 2.11

Table 3.2 - Relationship between the Sub Research Questions, Hypothesis and Literature Review

Sub Research Questions	**Hypothesis**	**Literature Review**
RQ 1.1	H1	2.9
RQ 1.2	H2	2.10
RQ 1.3	H3	2.11

3.6 Research Steps

Research steps are considered as the framework of the study, the same as the design structure of a house (Leavy, 2017). So, the explanation of the research steps is included in this section. The purpose of this research is to test the efficiency of the combination weightage of review helpfulness, review time, product rating and sentiment evaluation to design a rating system for the product in an e-commerce platform. Therefore, this research will conduct using the following steps. The graphical representation for the proposed research steps is shown in Figure 3.6.

Step 1: Data collection: This study will collect some real customer reviews which are publicly available from amazon.com.

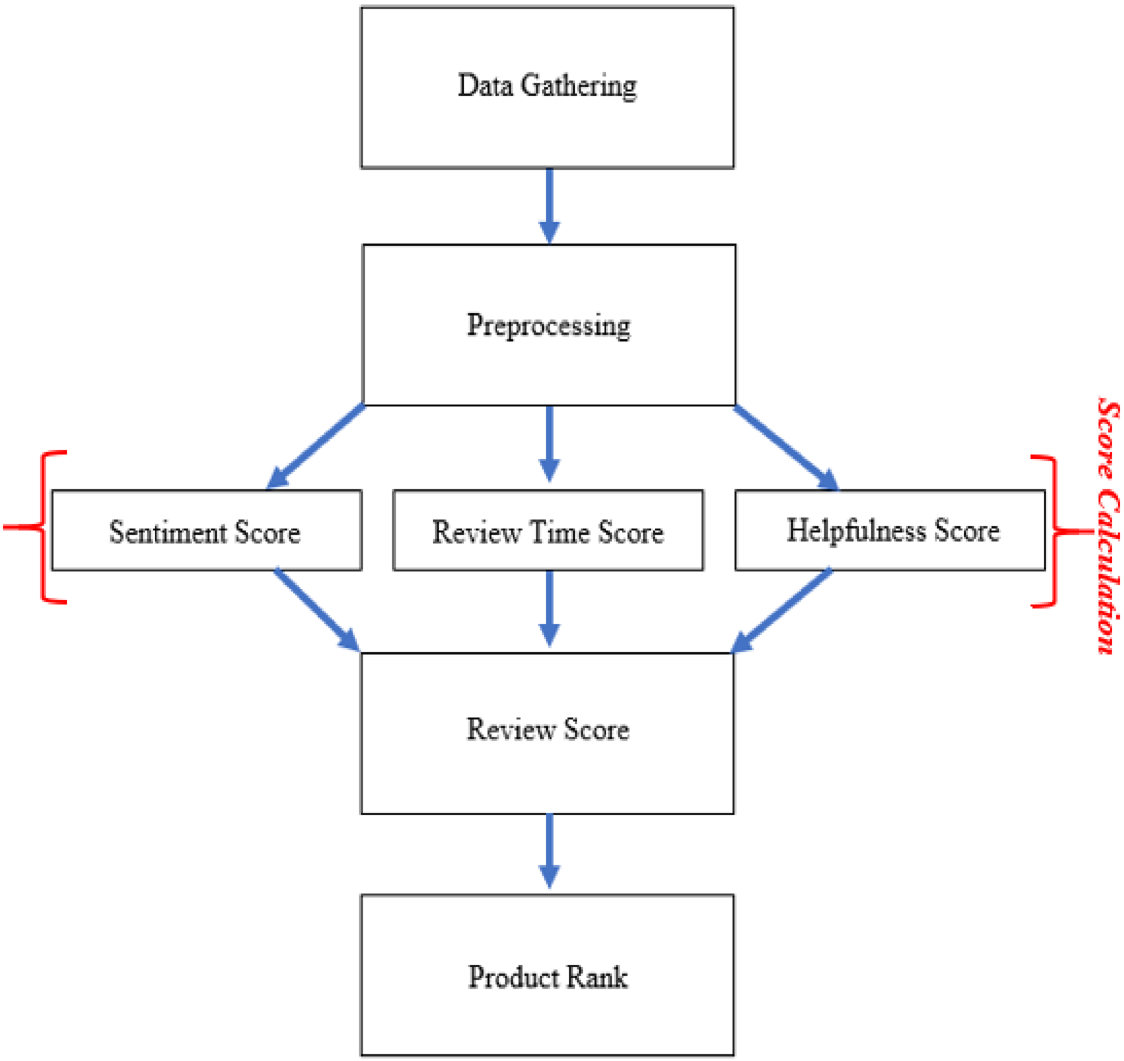
Step 2: Score calculation for helpfulness, relevancy (Time) and sentiment: This study will calculate the weightage score for review helpfulness, review time and calculate the sentiment score of the review.

Step 3: Review Score: This study will combine the individual segment score to calculate the entire review's score.

Step 4: Product rank generation: Based on the review's weightage, this study will regenerate a rating system for the product in the e-commerce platform.

Step 5: Evaluation of the result: the performance of the new method for an accurate product rating appraisal would be cross-checked with the current star rating system.

Figure 3.6 – Research Steps



3.7 Data Gathering

The bulk of earlier studies focussed on extracting the semantic features of customer review to analyse its rating, but this study will consider other factors to suggest a relevant rating to the review by considering review time and review helpfulness. The reviews from amazon.com, a well-reputed eCommerce platform, consist of review helpfulness and reviewed time along with the customer review. The Raptor web crawler (https://tools.raptor-dmt.com) is used to extract the reviews from amazon.com, and the majority of the data are available to the public from the free data set providers. Figure 3.8 depicts the appearance of a Raptor Web Crawler.

Figure 3.7 - Raptor Web Crawler

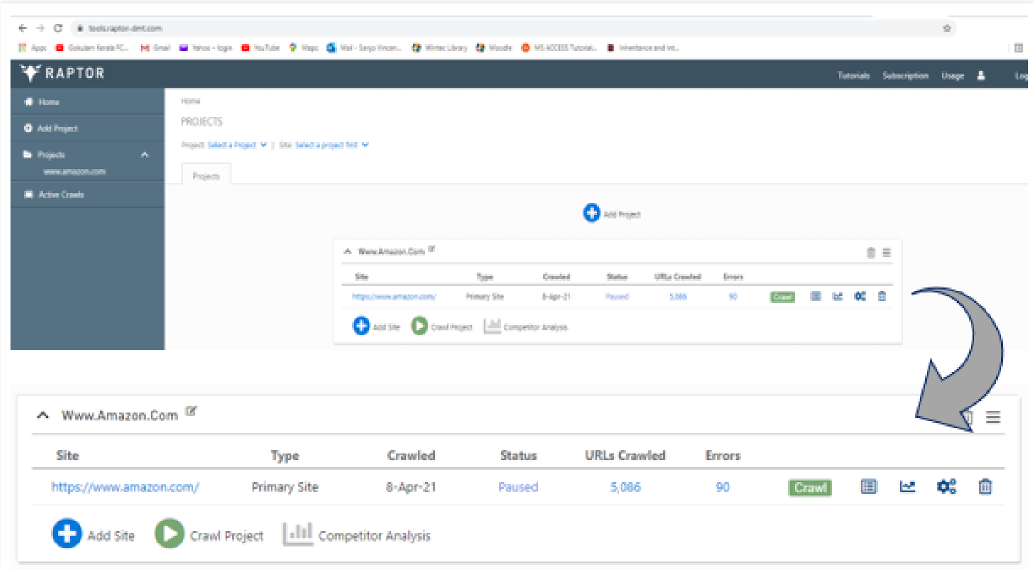
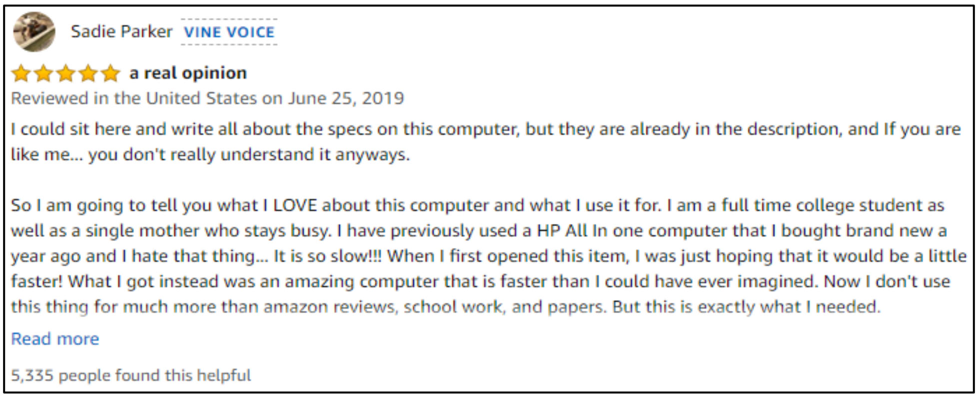


Figure 3.8 - Sample customer review from amazon.com


Sadie Parker VINE VOICE

a real opinion

Reviewed in the United States on June 25, 2019

I could sit here and write all about the specs on this computer, but they are already in the description, and If you are like me... you don't really understand it anyways.

So I am going to tell you what I LOVE about this computer and what I use it for. I am a full time college student as well as a single mother who stays busy. I have previously used a HP All In one computer that I bought brand new a year ago and I hate that thing... It is so slow!!! When I first opened this item, I was just hoping that it would be a little faster! What I got instead was an amazing computer that is faster than I could have ever imagined. Now I don't use this thing for much more than amazon reviews, school work, and papers. But this is exactly what I needed.

Read more

5,335 people found this helpful

Figure 3.9 shows an example customer review from Amazon.com, followed by a JSON representation of a customer review from Amazon.com.

Sample Review from amazon.com (JSON Format)

{ "reviewerID": "A2SUAM1J3GNN3B", "asin": "0000013714", "reviewerName": "J. McDonald", "helpful": [2, 3], "reviewText": "I bought this for my husband who plays the piano. He is having a wonderful time playing these old hymns. The music is at times hard to read because we think the book was published for singing more than playing from. Great purchase though!", "overall": 5.0, "summary": "Heavenly Highway Hymns", "unixReviewTime": 1252800000, "reviewTime": "09 13, 2009" }

Where reviewerID - ID of the reviewer, e.g. A2SUAM1J3GNN3B asin - ID of the product, e.g., 0000013714

reviewerName - name of the reviewer

helpful - helpfulness rating of the review, e.g., 2/3

reviewText - text of the review

overall - rating of the product

summary - summary of the review

unixReviewTime - time of the review (Unix time)

reviewTime - time of the review (raw)

3.8 Data pre-processing and evaluation

Upon completing the data collection, this analysis begins pre-processing using the BERT model, splitting the entire sentences into fragments of words and assigning some tokens to it for the next level processes like sentiment calculation. The next step is to conduct data analysis using an algorithm that the researcher uniquely develops to assign scores to the independent variables after data collection and data pre-processing. That algorithm, along with the sentiment score of the review, will measure the score for review time and review helpfulness. The average review score and a product rating based on the same standard would be given by combining individual scores.

3.9 Conclusion

This segment addressed the research methodology as well as the data collection for the study. The method of this study was the experiment method of quantitative research method. To combine variables and hypotheses, the researcher used an updated DSR model. Following section, Chapter 4 presents the scientific findings that were used to test the hypothesis.

4. ALGORITHMS

4.1 Introduction

Cormen (2009) informally defined that "an algorithm is any well-defined computational procedure that takes some value, or set of values, as input and produces some value, or set of values, as output". Algorithms are used to find the procedures, significant decisions and variables needed to address the problem. The development of an algorithm makes the solution mechanism rationally examined and even enforced (Knuth, 1997; Horowitz, 1997; Ogihara, 2018). Algorithms are often used to solve problems, so various algorithms were used in this analysis to properly solve the research goal. Chapter 4 delves into the various algorithms used in the analysis and discusses how they operate.

Section 4.2 explains the overview of the research process and section 4.3 depicts the algorithm to calculate the review helpfulness score of customer reviews or eWOM. Section 4.4 introduces the review time score calculation, which provides a relevant review time score for each review participated for the specific product. The review time relevance score calculation algorithm setting ten years as a realistic period for calculating the time score. One of the major sections of this study is outlined in Section 4.5, which holds the information about the review sentiment score calculation proceedings. Review sentiment score plays a vital role in overall review score calculation, which is explained in section 4.6. Section 4.7 illustrating the product rating calculations, and the conclusion is presented in Section 4.8 of this chapter.

4.2 Overview of the Research Process

This research is mainly divided into four steps, as described in Chapter 3, Section 3.6: Firstly, using web crawling software, the researcher gathers actual data from a website that specialises in collecting consumer feedback, such as amazon.com, and then pre-processes it. Second, the researcher gives numerical scores for helpfulness, time, and sentiment orientation for each analysis. Thirdly, based on the pre-calculated results, we compute a summary rating. Finally, the weighted average of the consumer review scores is used to calculate a product rating score.

4.3 Algorithm to Calculate eWOM Review Helpfulness Score

A study by Malik (2020) states that Hundreds of ratings for brands or products are continuously reported on e-commerce websites, so the number of feedbacks is exponentially growing, resulting in a situation of data overload. To address this problem, several e-commerce providers have implemented a feedback system that allows users to vote on a review which is helpful or not.

The quantity of support votes connected to an online customer review shows how communicative it is. The review which gets higher votes from different other users ordinarily gives more help to the others. As a result, since the helpfulness aspect is so important in this study, the researcher devises a method for calculating a score for the helpfulness of a review.

The following equation is used as the function for review the helpfulness score.

Equation 1 - Equation for review helpfulness score calculation

$$HS(Cr_{ij}) = \begin{cases} 0.75 & if\ Ch_{ij} = 0\ or\ \dfrac{\log_{10} Ch_{ij}}{Nh_j} \leq 0.75 \\ \log Nh_j(Ch_{ij}) & otherwise \end{cases}$$

Cr_{ij} : Review number. *i* represents the number and *j* represents the entity

$HS(cr_{ij})$: Helpfulness score for review cr_{ij}

Ch_{ij} : The number of votes got for the review Cr_{ij} for its helpfulness.

Nh_j : Total helpfulness votes for the most voted review for the entity *j*

The researcher would like to give a range of 0.75 to 1 as the review helpfulness score because the researcher believes that sometimes there are only a few helpfulness votes for entities. High helpfulness remarks earn a high score for helpfulness, and the poor helpfulness remarks also receive a low level of scores. Algorithm 1 explains how the researcher calculates the review helpfulness score for each customer review.

Algorithm 1: Score Calculation for eWOM Helpfulness

1: Define : $CRH_j = \{Ch_{1j}, Ch_{2j}, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots Ch_{nj}\}$

Input : CRH_j

Start Function $HS(Ch_{ij})$:

If $Ch_{ij} = 0$ or $\frac{\log_{10} Ch_{ij}}{Nh_j} \leq 0.75$ then

$HS \longleftarrow 0.75$

Else

$HS \longleftarrow \log_{max(CRH_j)}(Ch_{ij})$

End If

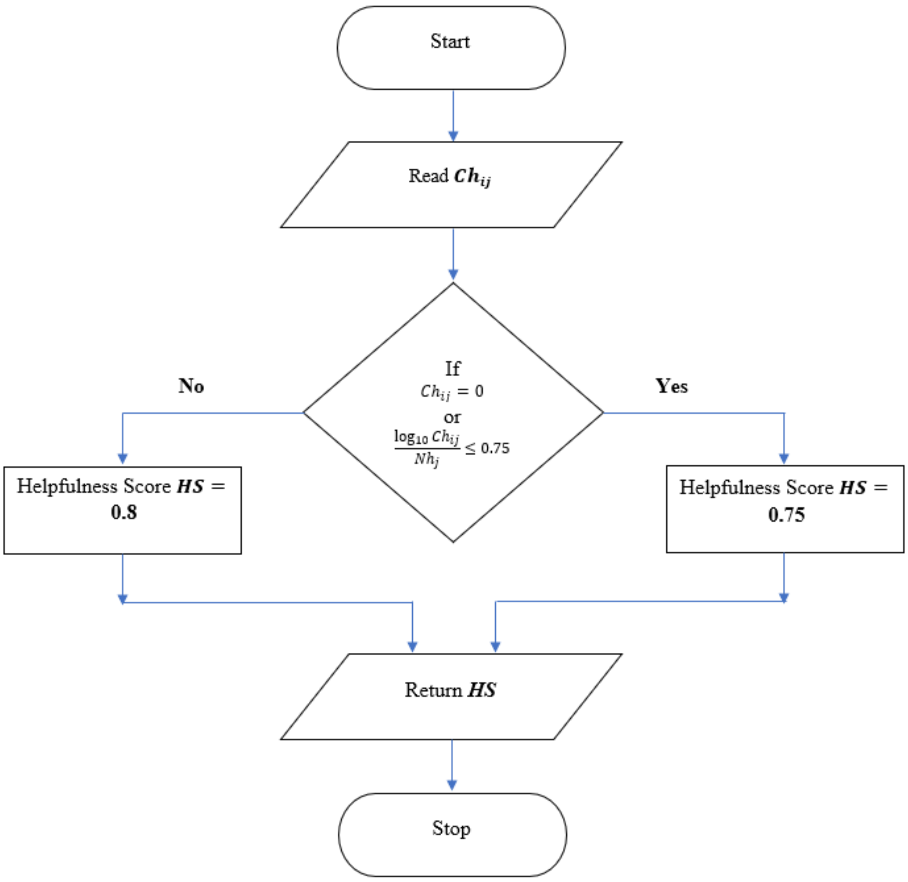
Return HS

End Function

CRH_j: Set of helpfulness ratings provided to a product

The following figure 4.1 depicts the flow diagram of the review helpfulness calculation function used in Algorithm 1.

Figure 4.1 - Flowchart of Review Helpfulness Score Calculation Function



4.4 Algorithm to Calculate time relevance score for an eWOM

Once a product, notebook, smartphone or even an iron box may get really good feedback, the years have taken away the product's strength, value and also influence consumers' judgments and decisions. Ultimately, "time is irreversible". For example, what will happen if we put a ten-year-old gaming PC with 100,000 (a hundred thousand) 5-star ratings in an online store? Nobody would think for a 10-year-old gaming PC, even though it was a best seller at the time with 100,000 5-star ratings. What is the reason for this? because new usage criteria have made it outdated. Its 100,000 reviews are no longer relevant. Reviews, like all those things and people, have an end date (Pownall, 2015). They become meaningless to the buyer at this stage. So, the researcher recommends that the review time's relevance should be considered a very valued factor for an online product rating system.

According to the researcher, the latest feedback help consumers with some more recent knowledge, so this study suggests the following equation to give each analysis a time score.

Equation 2 - Equation for Review Time Score Calculation

$$TS(Cr_{ij}) = \begin{cases} 0.8 & if\ Cy - Ry_{ij} \geq 10 \\ 1 - (Cy - Ry_{ij}) * 0.002 & otherwise \end{cases}$$

Cr_{ij} : Review number. *i* represents the number and *j* represents the entity

$TS(cr_{ij})$: Time score for review cr_{ij}

Ry_{ij} : Review time (Year) of review cr_{ij} .

Cy : Current Year

Algorithm 2: Score Calculation for Review Time

1: Define : $CRT_j = \{Ry_{1j}, Ry_{2j}, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots Ry_{nj}\}$

2: Input : CRT_j

3: Start Function $TS(Ry_{ij})$:

4: If $Cy - Ry_{ij} \geq 10$ then

5: $TS \longleftarrow 0.8$

6: Else

$TS \longleftarrow 1 - (Cy - Ry_{ij}) \times 0.002$

End If

Return TS

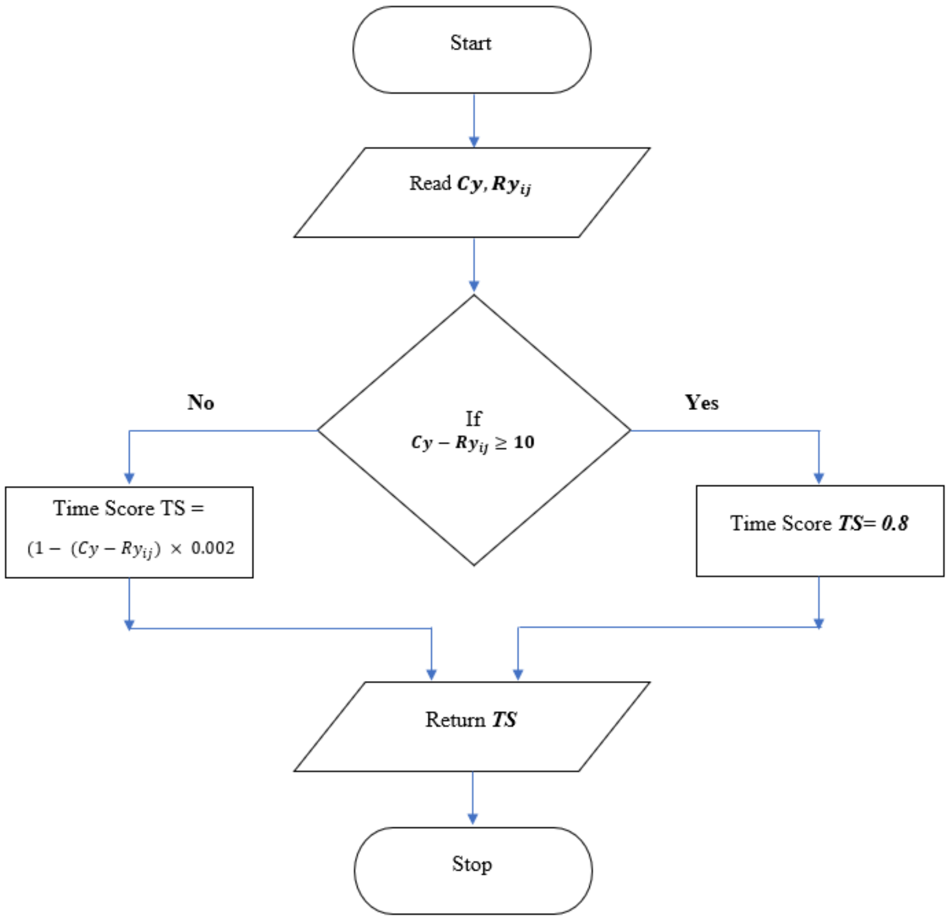
End Function

CRT_j: Set of review time provided to a product

The researcher used ten years as a realistic period for calculating the time score; if the period exceeded the period, the researcher used 0.8 as the review time score for each review. The researcher discovered that the rating score of 0.8 and >0.8 provided successful results to the studies with the aid of several other feedback. As a result, the researcher opted to use the same score for this study as well.

Figure 4.2 outlines the function for calculating the review time relevance score used in Algorithm 2.

Figure 4.2 - *Flowchart of Review Time Calculation Function*

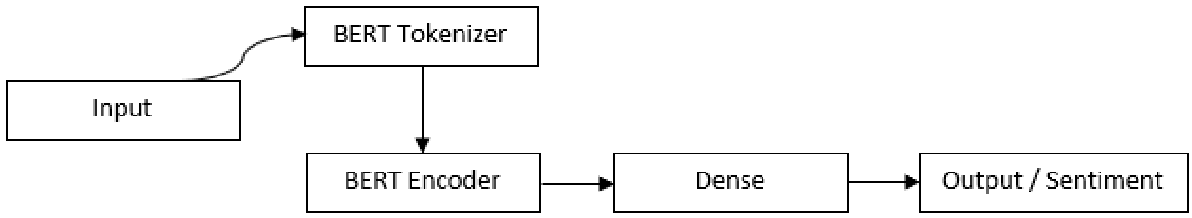


4.5 Algorithm to Calculate Sentiment Orientation of eWOM

The method of measuring positive or negative emotion in the text is known as sentiment analysis. Nowadays, companies use it to track sentiment in social media data, assess product quality, and better understand their clients (Puschmann & Powell, 2018; De, 2017; Santhipriya & Rao, 2018). Sentiment Analysis mainly focusing on sentiment polarities such as positive, negative, and neutral. Sentiment evaluation is important since it allows companies to easily consider their consumers' overall views (Jana & Uma, 2019; Sivasakthi, 2020). Algorithms that measure the tone of a transcript on a scale of positive to negative allow sentiment scoring.

BERT was developed to automate the encoding, analysis, and manipulation of language data such as speech and text using Natural Language Processing (NLP) techniques. BERT achieves this by merging two potent technologies: It is built on a profound network of transformers that effectively processes long texts with care. BERT is considered bi-directional. In other words, the entire text passage is taken into account to explain the meaning of each word. The model the researcher used for this analysis is outlined in Figure 4.3.

Figure 4.3 – Process flow of Sentiment Analysis using BERT (Devlin et al. (2019))



The researcher has trained the BERT model to assess the feeling orientation likelihood of specific analysis, and it helped to reach satisfactory outcomes by studying semantic relationships between terms or sub-words in a text in a wide range of NLP methods. This sentiment score generation algorithm assigns each analysis a score for its sentimental orientation. For this purpose, the researcher developed a formula as shown below and assigned the max() function to the output of the BERT Model.

Equation 3 - Equation for Sentiment Score Calculation

$$SS\left(Cr_{ij}\right) = max(BO\, neg._{ij}\, , BO\, pos._{ij}\,)$$

Cr_{ij} : Review number. *i* represents the number and *j* represents the entity

$SS\left(cr_{ij}\right)$: Sentiment score for review cr_{ij}

$BO\ neg._{ij}$: Negative sentimental orientation output from BERT

$BO\ pos._{ij}$: Positive sentimental orientation output from BERT

According to the researcher, Algorithm 3 is used to calculate the sentiment score using BERT finetuned model.

Algorithm 3: Score Calculation for Review Sentiment

1: Define : $CRS_j = \left\{Rs_{1j}, Rs_{2j}, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots Rs_{nj}\right\}$

$BO_j = \left\{BO(Rs_{1j}), BO(Rs_{2j}), \ldots\ldots\ldots\ldots\ldots BO(Rs_{nj})\right\}$

2: Input : CRS_j

3: Start Function $SS\left(Rs_{ij}\right)$:

4: If $BO\ neg._{ij} > \ BO\ pos._{ij}$ then

4: $SS \longleftarrow -1 \times$ *(max(*$BO(Rs_{ij})$*))*

6: Else

7: $SS \longleftarrow$ *max(*$BO(Rs_{ij})$*)*

5: Return SS

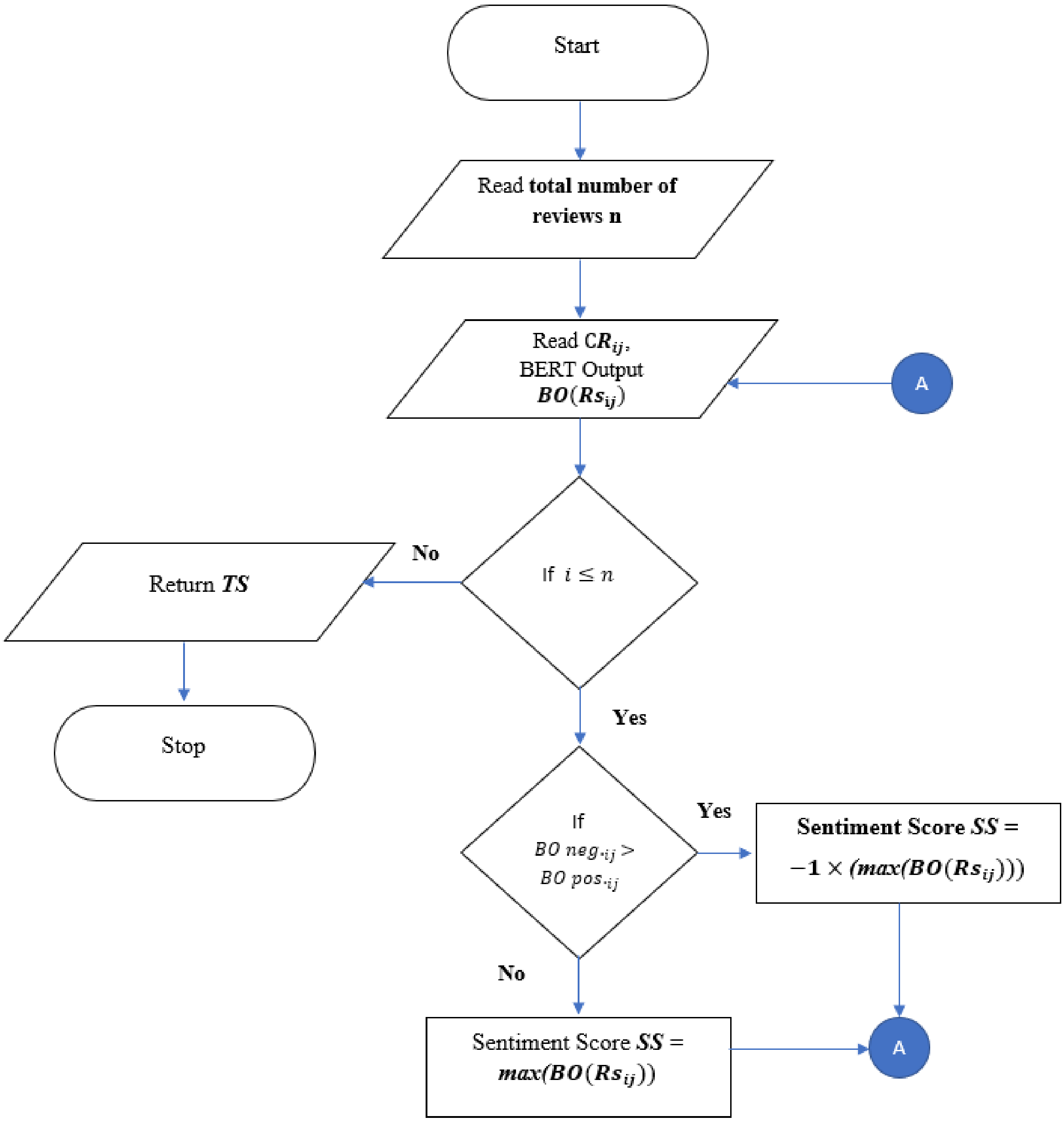
6: End Function

CRS_j: Set of reviews provided to a product

BO_j : Set of BERT model Output

If $BO\ neg._{ij} > \ BO\ pos._{ij}$ in the BERT measurement, it is considered a negative statement; otherwise, it is considered a positive statement. If the review is positive, it has a better chance of receiving a higher score than if it is negative. Figure 4.4 illustrating the Flow chart for Sentiment Score Calculation Function

Figure 4.4 - Flow chart for Sentiment Score Calculation Function



4.6 Algorithm to Calculate Overall Review Score

To support customer choice, it is necessary to provide a prospective client or consumer with enough detail or product rating based on users' previous experience. Thus, by presenting the generated review score to the target people, the researcher recommends analysing credibility. Overall review score calculation is the core part of automated product rating calculation. The overall review score redefines the review rating based on the previous scores such as review helpfulness (***HS***), review time (***TS***) and review sentiment orientation (***SS***). The overall score is considered as the average of HS, TS and SS. The following formula is used to calculate the overall review score for this study.

Equation 4 - Equation to Calculate Overall Review Score

$$ORS(Cr_{ij}) = \frac{HS(Cr_{ij}) + TS(Cr_{ij}) + SS(Cr_{ij})}{3}$$

Cr_{ij} : Review number. *i* represents the number and *j* represents the entity

$RS(cr_{ij})$: Review score for review cr_{ij}

$HS(cr_{ij})$: Review helpfulness score for review cr_{ij}

$TS(cr_{ij})$: Review time score for review cr_{ij}

$SS(cr_{ij})$: Review sentiment score for review cr_{ij}

Algorithm 4 details the overall review score calculation based on the equation above.

Algorithm 4: Overall Review Score Calculation for User Review

1: Define : $CR_j = \{Cr_{1j}, Cr_{2j}, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots Cr_{nj}\}$

2: Input : CR_j

3: Loop : for *i* in a *range(n)* do

$$ORS_{ij} = \frac{HS(Cr_{ij}) + TS(Cr_{ij}) + SS(Cr_{ij})}{3}$$

4: End Loop: End For

5: Output : ORS_j

ORS_{ij} : Review score for review *i* for the product *j*

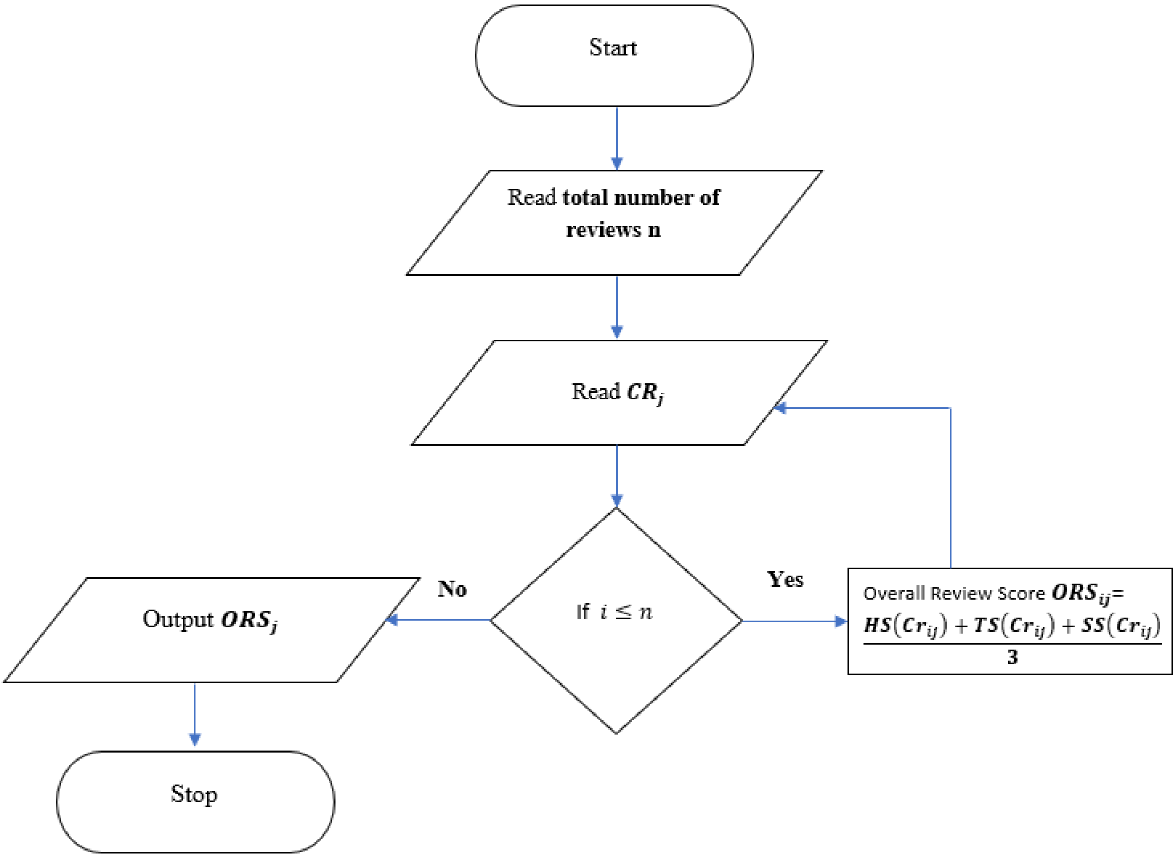
CR_j: Set of reviews provided to a product *j*

Cr_{ij} : Review number. *i* represents the number and *j* represents the entity

n: Number of reviews.

Since all of the other ratings, such as HS, TS, and SS, are between 0 and 1, the RS must be between 0 and 1. The flow chart for calculating review scores is depicted in Figure 4.5.

Figure 4.5 - Flow chart for Overall Review Score Calculation Function



Example 1: A review has a negative sentiment orientation and a review helpfulness score of 0.78, as well as scores of 0.97 and -0.99 for review time and sentiment, the review score is measured as follows.

$$\textbf{Review Score (ORS)} = \frac{\mathbf{0.78+0.97-0.99}}{\mathbf{3}} = \mathbf{0.25}$$

Example 2: A review has a positive sentiment orientation and a review helpfulness score of 1, as well as scores of 0.96 and 0.98 for review time and sentiment, the review score is measured as follows.

$$\textbf{Review Score (ORS)} = \frac{\mathbf{1+0.96+0.98}}{\mathbf{3}} = \mathbf{0.98}$$

Aside from the review time and helpfulness, the overall review score is influenced by the sentiment score. If there are more negative feedback than positive reviews, the review score will be influenced by the negative reviews, and the review score will be held at a minimum value, as in Example 2.

4.7 Algorithm to Calculate Review Rating

Customers usually use the star rating mechanism to rate products via eCommerce platforms. The reviews and product ratings are critical in establishing the brand's credibility because they show the upcoming consumers that the product is highly valued and trusted by prior customers at a glance. However, since star ratings are solely based on the customer's perspective, the ratings relating to the reviews can often confuse readers, such as where a poor review can be misinterpreted as an outstanding 5-star score. One of the major eCommerce platforms Amazon.com, following machine learning algorithms to calculate its product star rating variables like the date of purchase and user authenticity by verifying whether the product originally purchased or not (Amazon.com, n.d.). In this report, the researcher also uses an additional Amazon feature called "amazon verified purchase," which ensures Amazon has verified that the individual writing the user review bought the product from Amazon and did not buy it at a significant discount (Amazon.com, n.d.).

Customer buying habits are heavily influenced by product reviews, particularly on eCommerce platforms. When a customer hovers over a product picture or data, they want immediate access to its overall rating to analyse the product is fit for them. These kinds of access will help them see the total number of purchased people and the number of reviewed customers and their reviews scores. The researcher believing that this new product rating system will consider a 365-degree view of the product reviews and suggesting an appropriate rank for the particular product entirely based on the customer reviews.

Based on the overall review ratings for the consumer feedback engaged by the consumers, the following equation is used to measure the Review Rating and Algorithm 5 details the overall review score calculation based on the Equation 5.

Equation 5 - Equation to Calculate Review Rating

$$PR_j = \left(\frac{ORS_{1j} + ORS_{2j} + ORS_{3j} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots ORS_{nj}}{n}\right) \times 100$$

ORS_{ij} : Review score for review *i* for the product *j*

PR_j: Review Rating score provided to a product *j*

Algorithm 5: Overall Review Score Calculation for User Review

1: Define : $CR_j = \{Cr_{1j}, Cr_{2j}, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots Cr_{nj}\}$

2: Input : CR_j

3: *tempORSSum* $\longleftarrow$ 0

4: Loop : for *i* in a *range(n)* do

$tempORSSum = tempORSSum + ORS(Cr_{ij})$

5: End Loop: End For

6: $PR_j \longleftarrow \left(\frac{tempORSSum}{n}\right) \times 100$

5: Output : PR_j

$ORS(Cr_{ij})$: Overall Review score for review *i* for the product *j*

Cr_{ij} : Review number. *i* represents the number and *j* represents the entity

PR_j: Review Rating score provided to a product *j*

n: Total Number of reviews.

For example, Table 2 shows how Review Ratings are calculated. To illustrate the Review Rating calculation, this description would only consider 7 reviews of a single product. The scores they received in each process are listed in Table 4.1. In addition to that, Table 4.2 shows how the total review score and star rating for this study are linked. The flowchart for Algorithm 5 is depicted in Figure 4.6.

Figure 4.6 - *Flow chart for Calculation of Review Rating*

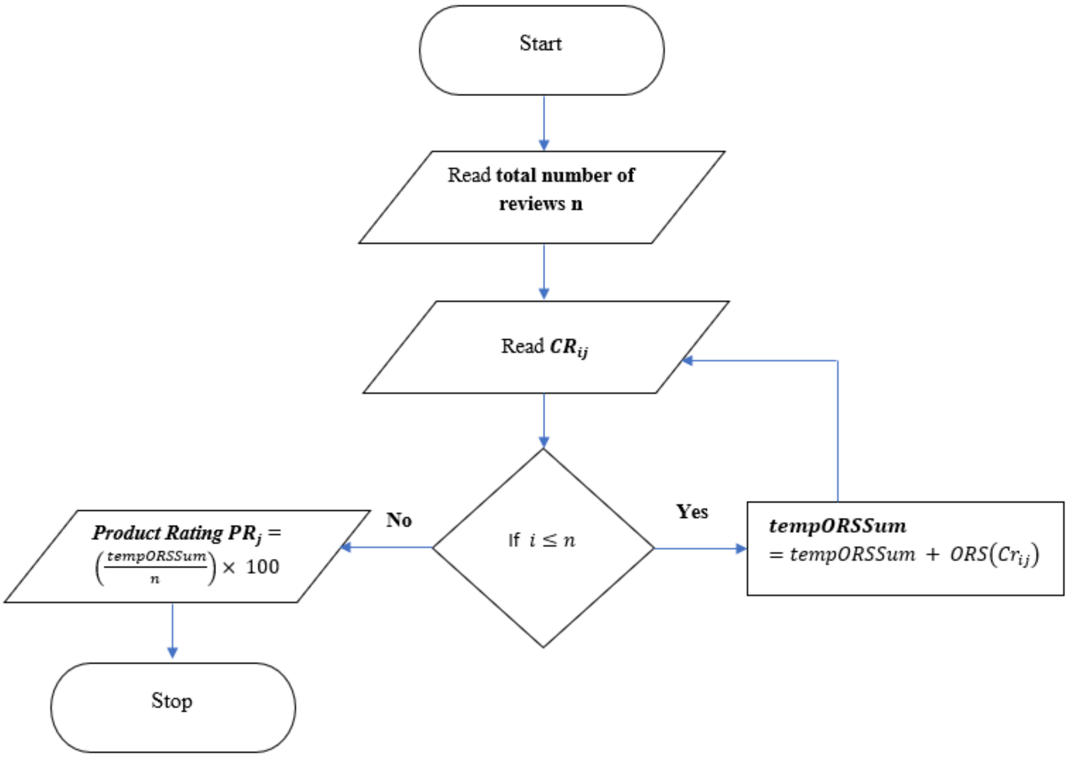


Table 4.1 - Example for Review Rating Calculation

<table>
<tr><td rowspan="2">Review</td><td colspan="6">Review Scores</td><td rowspan="2">Rating Score</td><td rowspan="2">Review Rating</td></tr>
<tr><td>Helpfulness (HS)</td><td>Time (TS)</td><td>Sentiment Orientation</td><td>Sentiment (SS)</td><td>Overall (ORS)</td><td>Sum (ORS)</td></tr>
<tr><td>Rw1</td><td>0.78</td><td>0.96</td><td>Negative</td><td>0.99</td><td>0.25</td><td rowspan="7">0.72</td><td rowspan="7">72</td><td rowspan="7">4 Stars</td></tr>
<tr><td>Rw2</td><td>1</td><td>0.95</td><td>Positive</td><td>0.88</td><td>0.94</td></tr>
<tr><td>Rw3</td><td>0.75</td><td>0.80</td><td>Positive</td><td>0.93</td><td>0.82</td></tr>
<tr><td>Rw4</td><td>1</td><td>0.83</td><td>Negative</td><td>0.96</td><td>0.29</td></tr>
<tr><td>Rw5</td><td>0.86</td><td>0.92</td><td>Positive</td><td>0.96</td><td>0.91</td></tr>
<tr><td>Rw6</td><td>0.91</td><td>0.85</td><td>Positive</td><td>0.92</td><td>0.89</td></tr>
<tr><td>Rw7</td><td>0.94</td><td>0.92</td><td>Positive</td><td>0.97</td><td>0.94</td></tr>
</table>

Table 4.2 - Relationship between total review score and star rating

Review Rating Score	Corresponding **'Star Rating'**
In-between **0** and **10**	**N.A**
In-between **10** and **30**	**1 Star**
In-between **30** and **50**	**2 Stars**
In-between **50** and **70**	**3 Stars**
In-between **70** and **90**	**4 Stars**
In-between **90** and **100**	**5 Stars**

4.8 Conclusion

To conclude, the eCommerce service provider will show the consumer the product's total star rating and the number of feedbacks based on the previously mentioned algorithms. Aside from that, new and positive reviews have a greater influence on the product ranking score than old and unhelpful reviews. The following section, Section 5 explains the analysis and results of this particular study.

5. ANALYSIS

5.1 Introduction

This chapter aims to examine the user review using the data collection strategy that was chosen. In a great manner, the analysis is planned and organized and this approach is used to analyse the data in this overall research process. Section 5.2 describes the data pre-processing and section 5.3 explains the BERT fine-tuning. In section 5.4 produce a detailed analysis of the sentiment analysis process. Finally, Section 5.5 describes how the algorithms are executed and section 5.6 concludes this chapter with a summary of this chapter.

5.2 Data Pre-processing

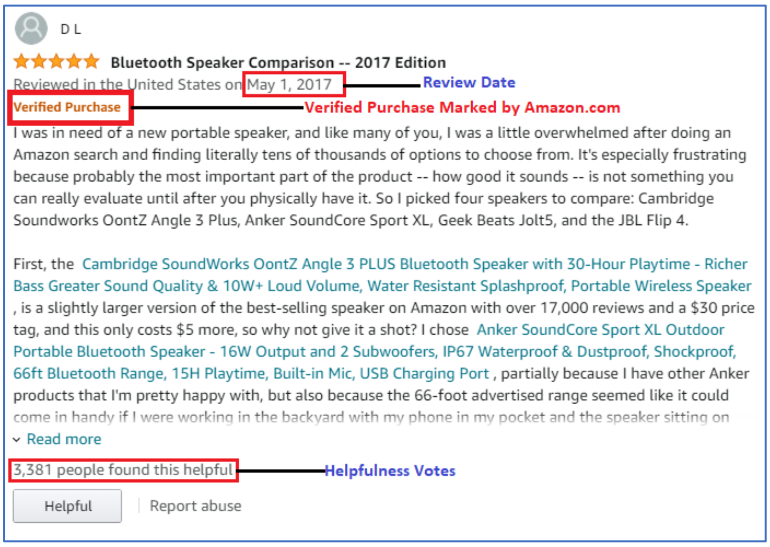
Customer reviews from Amazon.com for four separate product segments (especially product manufacturers) were the subject of this study. Using the scaping app Raptor, the researcher gathered user feedback from Amazon. We gathered 600 reviews for three different smartphone brands, 300 reviews for two different notebook brands, 400 reviews for four different action camera brands, and 200 reviews for two different speaker brands. Since all of the ratings are publicly accessible and Amazon.com has already checked and marked as a verified order, the researcher believes that the results or customer reviews are more reliable. Table 5.1 describes the categorisation of each product categories used in this research.

Table 5.1 - No. of Reviews used in this study.

Product Category	Total No. of Reviews Extracted	No. of Products (in each category)	Reviews per Product	Amazon.com Verified Purchase
Smart Phones	600	3	200	Verified
Laptops	300	2	150	Verified
Action Camera	400	2	200	Verified
Speakers	200	2	100	Verified

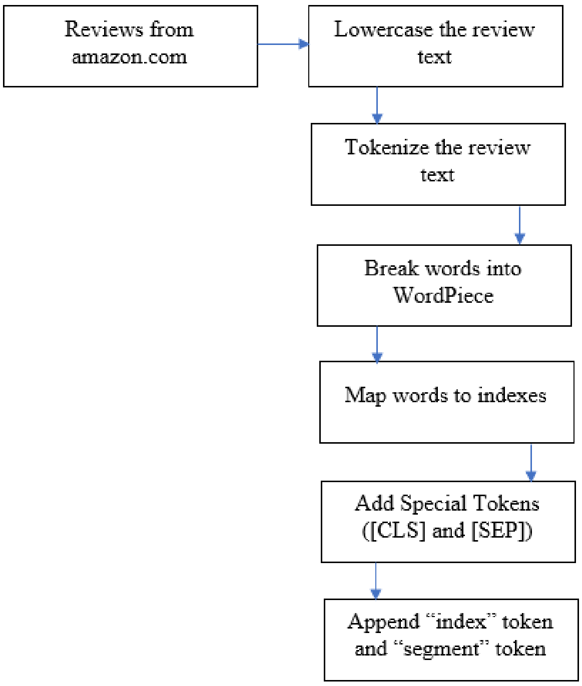
Any extracted review includes the reviewer's helpfulness votes, the exact review time, and the review message. Figure 5.1 depicts an amazon.com review used in this study.

Figure 5.1 - Model review from amazon.com used in this research



After gathering the reviews from Amazon.com, the researcher converts the entire text to lower case to adopt the BERT lowercase model, tokenises the text, and breaks each word into WordPiece, a sub word segmentation algorithm used in NLP. Next is the mapping process. To complete the BERT evaluation the researcher, need to map the words to indexes as per the file provided by BERT. In addition to this, the researcher needs to add two special tokens, "CLS" and "SEP" along with that. Finally, each input should append "index" and "segment" tokens. Figure 5.2 illustrating the pre-processing procedure for the customer reviews.

Figure 5.2 - Pre-processing flow of Customer Reviews fetched from amazon.com



The next section will shed some light on different score calculations along with the BERT model sentiment analysis.

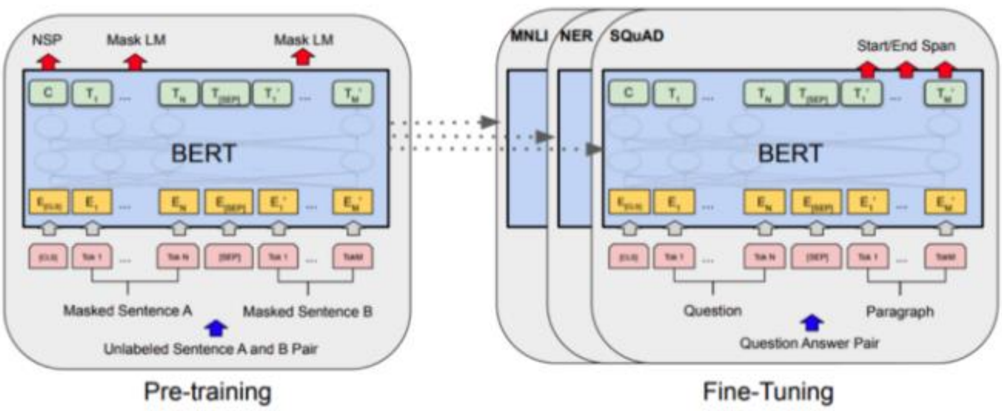
5.3 BERT Fine-Tuning for Sentiment Analysis

The introduction of BERT, which has been described as the start of a new age in NLP, is among the most significant landmarks in the progress of NLP lately. This study uses the HuggingFace[1] transformers library to fine-tune the BERT model. The efficiency of the BERT base and the fine-tuned one will then be compared using a TF-IDF vectorizer and a Naive Bayes classifier. According to the researcher, the transformer library helps him fine-tune the new BERT models quickly and comfortably, achieving a 10 per cent higher accuracy rate than the base model. Figure 5.3 illustrating the overall view of BERT fine-tuning. There are two types of pre-trained BERT models.

- BERT base with 12-layer,768=hidden,12-heads,110M parameter neural network.
- BERT large with 24-layer, 1024-hidden, 16-heads, 340M parameter neural network architecture.

This study fine-tunes the BERTbase model in this study to arrive at our sentiment score estimate for customer feedback retrieved from amazon.com.

Figure 5.3 - Pre-Training and Fine-Tuning Techniques for BERT in General (Devlin et al., 2019)



The sections that follow explain how this researcher fine-tuned the BERT for sentiment analysis to measure the sentiment score of each review, specifically for this study.

[1] https://huggingface.co

5.3.1 Download Dataset and Load Train Data

The next important step in the BERT sentiment analysis process is to download the required data set. This study using the customer reviews from amazon.com for four different categories of products. The research loads the train data and labels it for future work. The training dataset should avoid unwanted columns and only keeping the mandatory fields such as customer id, product id, customer review etc. Figure 5.4 depicts the model dataset have used in this study.

Figure 5.4 - Sample Dataset

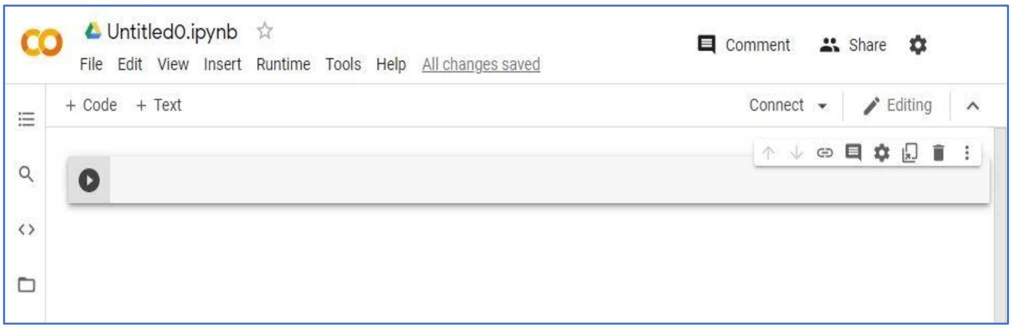
reviewerID	asin	reviewerName	helpful	reviewText	overall	summary	reviewTime
A2IBPI20UZIR0U	1384719342	cassandra tu "Y	0	Not much to write about h	5	good	02 28, 2019
A14VAT5EAX3D9S	1384719342	Jake	13	The product does exactly	5	Jake	03 16, 2019
A195EZSQDW3E21	1384719342	Rick Bennette "	1	The primary job of this de	5	It Does The Job	08 28, 2019
A2C00NNG1ZQQG2	1384719342	RustyBill "Sunda	0	Nice windscreen protects	5	GOOD WINDSCR	02 14, 2019
A94QU4C90B1AX	1384719342	SEAN MASLANK	0	This pop filter is great. It	5	No more pops w	02 21, 2019
A2A039TZMZHH9Y	B00004Y2UT	Bill Lewey "blew	0	So good that I bought ano	5	The Best Cable	12 21, 2019
A1UPZM995ZAH90	B00004Y2UT	Brian	0	I have used monster cable	5	Monster Standa	01 19, 2019
AJNFQI3YR6XJ5	B00004Y2UT	Fender Guy "Ric	114	I now use this cable to rur	3	Didn't fit my 19	11 16, 2019
A3M1PLEYNDEYO8	B00004Y2UT	G. Thomas "Ton	0	Perfect for my Epiphone S	5	Great cable	07 6, 2008
AMNTZU1YQN1TH	B00004Y2UT	Kurt Robair	0	Monster makes the best c	5	Best Instrument	01 8, 2019
A2NYK9KWFMJV4Y	B00004Y2UT	Mike Tarrani "Ja	6	Monster makes a wide arr	5	One of the best	04 19, 2019
A35QFQI0M46LWO	B00005ML71	Christopher C	75	I got it to have it if I neede	4	It works great bu	04 22, 2019
A2NIT6BKW11XJQ	B00005ML71	Jai	0	If you are not use to using	3	HAS TO GET USE	11 17, 2019
A1C0O09LOLVI39	B00005ML71	Michael	0	I love it, I used this for my	5	awesome	06 16, 2019

To obtain an expert result, a training data set with 90% of the data and a validation set with 10% of the data is divided by random division of the whole training data into two. The validation set will be used to compare the models while the train set is used for hyperparameter tuning through cross-validation.

5.3.2 GPU Setup

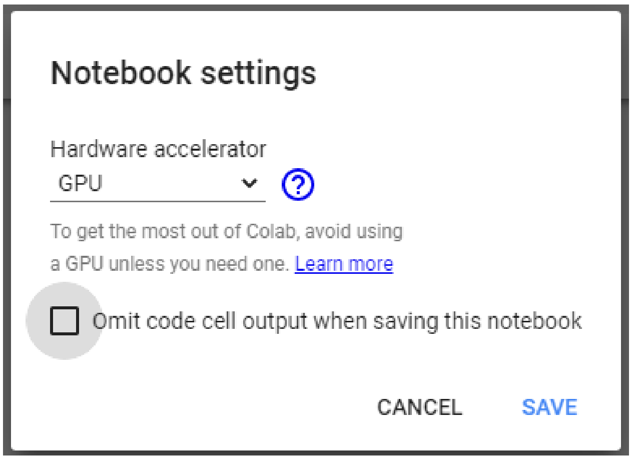
Google Colab provides GPUs and TPUs for free. It's better to use these features because this study will be training a neural network. Figure 5.5 depicts the interface of Google Colab.

Figure 5.5 - Google Colab Interface



To add a GPU in Google Colab, go to the "Runtime" menu item then select "Change runtime type" from the submenus. Then the Google Colab interface will show a window like the following Figure 5.6. Select "GPU" for the hardware accelerator and click on save for future use.

Figure 5.6 - Google Colab Notebook Settings



The GPU must now be defined as a device. Run the following code in the Google Colab interface to get a confirmation response similar to Figure 5.7.

Figure 5.7 - Google Colab GPU Status

```
import torch

if torch.cuda.is_available():
    device = torch.device("cuda")
    print(f'There are {torch.cuda.device_count()} GPU(s) available.')
    print('Device name:', torch.cuda.get_device_name(0))

else:
    print('No GPU available, using the CPU instead.')
    device = torch.device("cpu")
```

```
There are 1 GPU(s) available.
Device name: Tesla T4
```

Now the Google Colab GPU is ready to use.

5.3.3 Data Pre-processing

Bag-of-Words (BoW)[2] model interpreted the text as a bag of its words and ignoring the grammar and order of the word (Zhang et al., 2010; Verberne et al., 2010; Lei Wu et al.,

[2] Bag-of-words(BoW) model is a text representation model that represents the frequency at which words appear in a text. The model only cares about whether the words appear in the text, not where they appear.

2010). So, the researcher needs to avoid the words used to stop, the punctuations, and the characters that never affect the customer review's overall content. Until feeding the user review text data to machine learning algorithms, the term frequency-inverse document frequency (TF-IDF)[3] (Wu et al., 2008; Savyanavar, & Mehta, 2016; Springer US, 2017) would be used to vectorize it.

5.3.4 Tuning of Hyper Parameter

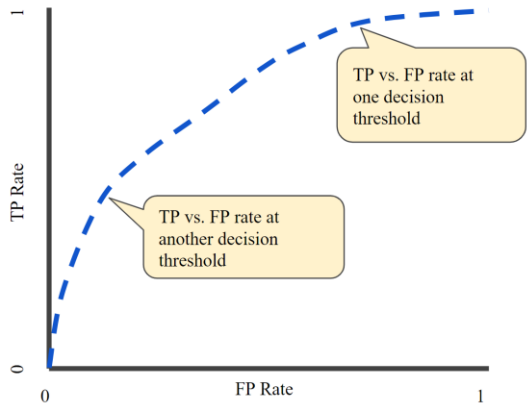
Tuning of Hyperparameter is one of the most important steps in BERT fine-tuning for sentiment analysis. Before explaining the hyperparameter tuning, we need to understand True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), Receiver Operating Characteristic (ROC) curve and area under the ROC Curve (AUC).

- ***True Positive (TP)*** - A true positive is a consequence where the positive class is accurately estimated by the model.
- ***True Negative (TN)*** - A true negative is a result where the negative class is accurately estimated by the model.
- ***False Positive (FP)*** - A false positive is an outcome where the positive category is falsely estimated by the algorithm.
- ***False Negative (FN)*** - A false negative is a result in which the negative class is falsely predicted by the model.
- ***Receiver Operating Characteristic (ROC) Curve*** - The ROC curve is a graph that depicts a classification model's effectiveness overall categorisation levels. True Positive Rate (TPR) vs. False Positive Rate (FPR) at various classification levels is plotted on a ROC curve. The reduction of the rating threshold categorises more products as positive and thus increases all false and true positives. A standard ROC curve is depicted in Figure 5.8. TPR and FPR are calculated by using the formulas:-

$$\text{TPR}= \frac{TP}{TP+FN} \qquad \qquad \text{FPR}= \frac{FP}{FP+TN}$$

[3] http://www.tfidf.com

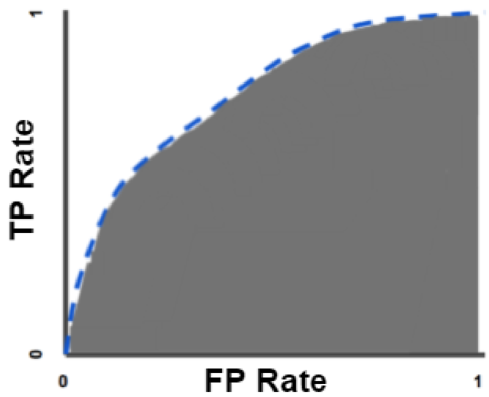
Figure 5.8 - TP vs. FP rate at different classification thresholds



Note. Google (n.d).*TP vs. FP rate at different classification thresholds.* (https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc)

- ***The area under the ROC Curve (AUC)* -** AUC measuring the entire area (0,0) to (1,1) of a ROC curve. AUC quantifying the performance that takes into account all potential classification thresholds. Figure 5.9 depicts an AUC.

Figure 5.9 - AUC (Area under the ROC Curve)

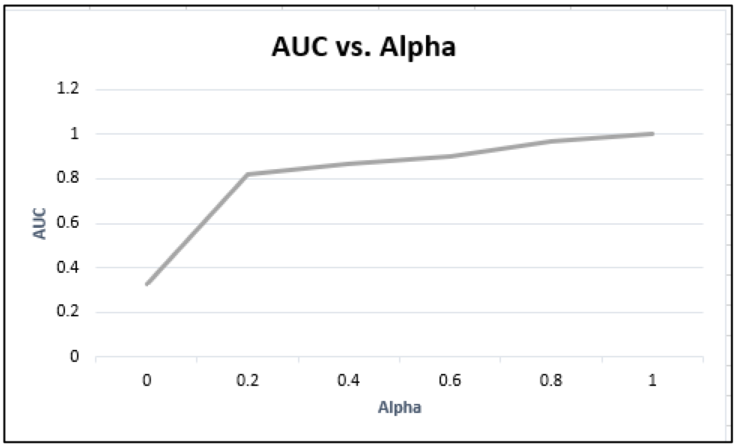


Note. Google (n.d).*TP vs. FP rate at different classification thresholds.* (https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc)

The hyperparameters will be fine-tuned using cross-validation and the AUC ranking. The function '*get_auc_CV*' returns the average score for AUC from cross-validation. To define discrete features, the Multinomial naive Bayes classifier is used. In this study, the MultinominalNB class holding only one hyperparameter called alpha. The Alpha value 1.3

provides us with the maximum AUC Score. Figure 5.10 showing the line diagram for the comparison between AUC and Alpha.

Figure 5.10 - AUC vs. Alpha Graph



5.3.5 Performance Evaluation of the Base Model

This study will measure the accuracy rate and the AUC score of the model used for this research to assess its results on the validation collection. On the validation range, this model obtains a 76.25 percent accuracy rate along with 0.8451 as AUC by integrating TF-IDF and the Naive Bayes algorithm. This is the starting point for evaluating the efficiency of our fine-tuned BERT model.

5.3.6 Installing Hugging Face Transformers Library

As per the Hugging Face website "Hugging Face Transformers (formerly known as PyTorch-transformers and PyTorch-pretrained-bert) provides general-purpose architectures (BERT, GPT-2, RoBERTa, XLM, DistilBert, XLNet) for Natural Language Understanding (NLU) and Natural Language Generation (NLG) with over 32+ pre-trained models in 100+ languages and deep interoperability between TensorFlow 2.0 and PyTorch"(huggingface.co, n.d). As per the researcher's view, the Hugging Face library seems to be the most commonly used and efficient Python framework for interacting with BERT at the present. The library supports a range of pre-trained transformer models and provides pre-built adaptations to these models that are tailored to this particular mission. The transformer library installation is comparatively easy by running a pip line with Google Colab.

5.3.7 Pre-processing of Text Data

Firstly, the text needs to be slightly processed to delete special characters interpreted as text and replace them with the right symbol (for example, shift '& amp;' to '&'). Since BERT was taught for the full statement, the level of processing would be much lower than in previous approaches.

For example,

Original: I'm having difficulties. I purchased this product last month, but I cannot log in & use it. Will you able to assist?

Processed: I'm having difficulties. I purchased this product last month, but I am unable to log in & use it. Will you able to assist?

In the mentioned example, the original text contains the '&' special character instead of the '&' sign then the text pre-processing methods will replace the special character text with the original character.

5.3.8 Tokenization

This experiment would have to use the tokenizer offered by the library to implement the pre-trained BERT. This is due to: -

(1) There has a defined dictionary in the BERT model

(2) There has a specific guideline to use the words out of the dictionary by the BERT tokenizer.

In addition, special tokens must be used in the beginning and finishing of each sentence, all statements must be truncated to one constant span, and the padding tokens with "attention mask" clearly specified. The BERT model allows us to use the *encode_plus* method and it will:

(1) make tokens out of our text,

(2) add [CLS] and [SEP] tokens,

(3) Convert tokens tokenizer vocabulary indices,

(4) Pad or truncate statements to specified pay-out, and

(5) Implement an attention mask.

Apart from this, before performing tokenization to specify the maximum length of the sentences plays an avital role in this overall process. After tokenizing the data, we get an output as shown in Figure 5.11.

Figure 5.11 - Tokenization Output

```
Original:  I'm having difficulties. I purchased this product last month, but I am unable to log in & use it.
Will you able to assist?
Token IDs:  [1045, 1005, 1049, 2383, 3314, 1012, 7483, 1045, 2128, 8654, 2098, 2005, 2484, 2847, 2044, 1045, 2001, 4011,
2000, 4875, 1010, 2085, 1045, 2064, 1005, 1056, 8833, 2006, 1004, 4638, 1999, 1012, 2064, 2017, 2393, 1029, 102, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
```

The use of the PyTorch[4] data loader during training can help conserve memory and increase training speed.

5.3.9 BERT Classifier and Validation Set Evaluation

BERT-base contains 12 transformer layers then each layer also produces the same number of token embeddings and holds its list. The [CLS] token is used to provide a classifier to the output of the final transformer layer. The transformer library includes a class called *BertForSequenceClassification* class is structured to classify the input. The researcher will however construct a new class so that he can use classifiers for himself. This study would establish an optimiser to fine-tune the BERT classifier with the hyper-parameters such as 16 or 32 batch size, 5e-5, 3e-5 or 2e-5 learning rate and 2,3 and 4 as several epochs. Huggingface contributed the run glue.py script, which shows how to use the transformers library. The AdamW optimiser is used in the experiment. After this, the researcher starts to train the loop and evaluates its results on the validation sets, then it is ready to start the BERT Classifier training.

The researcher will run a series of codes to compute logits in a forward pass and will apply SoftMax to quantify probabilities. Then the BERT Classifier achieves 0.90 AUC and 88.81 accuracies. Finally, train the model with the entire data. Furthermore, even though BERT is massive, complex, and has millions of parameters, need to fine-tune it in 2-4 epochs. BERT is one of the strongest NLP models available right now. Since BERT was educated on a large volume of data and has already encoded a lot of knowledge about our language, fine-tuning is slightly easier with these small datasets.

5.4 Sentiment Analysis

The researcher fine-tuned the BERT Base model to determine the sentiment preference of received feedback. The researcher developed the model by constructing a single new layer that was practised with a dataset that included both positive and negative processed

[4] https://pytorch.org

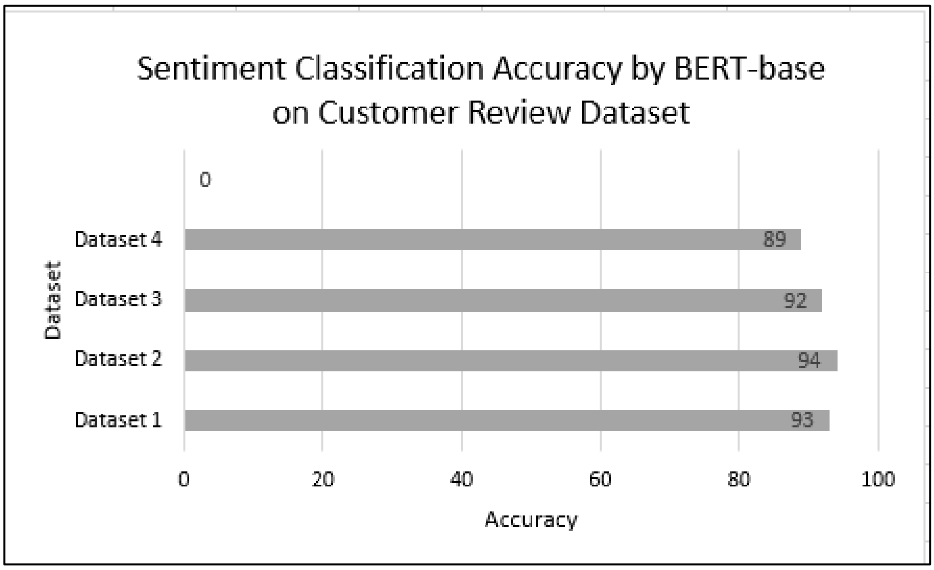
feedback. Table 5.2 explains the details of sequence length, batch size, learning rate and the number of epochs for this study.

Table 5.2 - Fine-Tuned BERT-base Model Settings

Sequence Length	128
Batch Size	32
Learning Rate	0.00002
Number of epochs	3

After experimenting with the BERT-Base model to test the accuracy of the received feedback, Figure 5.12 depicts the model's precision in sentiment orientation predicting.

Figure 5.12 - Sentiment Classification Accuracy on Customer Review Dataset



5.5 Experiment with Algorithms

According to the researcher, many aspects of Python make it an excellent platform for an NLP project. This language's easy syntax and precise semantics make it an ideal tool for Natural Language Processing projects. Furthermore, programmers will benefit from good convergence with other languages and technologies useful for machine learning techniques. It gives developers access to an extensive range of NLP resources and libraries, allowing them to perform tasks like document classification, topic modelling, POS tagging, word vectors, and sentiment analysis (Trappenberg, 2019; Silaparasetty, 2020; Stancin & Jovic, 2019).

To experience the output of this proposed model 'Samthripthi', the researcher develops a software program using the programming language Python (Version: 3.9.2, windows operating system release) and the researcher includes all algorithms mentioned in chapter 4 along with fine-tuned BERT model and found the results discuss in chapter 6. The Review ID and the corresponding text (user review) used to demonstrate the sentiment orientation score is shown in

Table 5.3. Using these randomly selected user reviews to calculate the review score based on the proposed algorithms by this study and all the results are presented in chapter 6. Chapter 7 discuss and interpreting these results.

Table 5.3 - Review ID and Review Text Using for Different Score Calculation

Review ID	**Review Text**	**Review Helpfulness Votes**	**Review Time**
A2A039TZMZHH9Y	"I paid it was worth the trouble of finding all this out, hope this helps you when it comes to spending your money and time. All because of the double SIM card slot you can add a business SIM card, this will give you to phone lines, personal and business."	13	1/12/2020
A195EZSQDW3E21	"Slower than I expected. Otherwise, I like the features."	28	13/11/2020
A3M1PLEYNDEYO8	"I wanted this phone so bad and I could not activate it. This is an international	159	05/02/2021

	version, so, if you are in the USA even though it says is unlocked won't work with any company. I am about to return it.”		
A35QFQI0M46LWO	“The product arrived on time but when I opened the box there was a European charger and a converter for us plugs. I picked up the phone and for some reason, I just did not like it so I returned it the next day. Unlike most things I buy on Amazon, it cost me money to return it and that I did not like.”	0	14/12/2020
A1GMWTGXW682GB	“The phone specs are inappropriately stated, it supposed to be a dual sim and support the international version of it but I have received a phone which has one sd card slot and a SIM card slot. Where is the other dual SIM card slot, this is so ridiculous, I bought it from Amazon as they are selling dual sim Samsung phone. This may support gsm phones; I don’t care where is my dual sim slot. I want this	4	30/11/2020

	to be fixed/ replaced / ASAP”		
A2FZ4Z0UFA1OR8	“I think I made a very wise decision when I bought this phone. with the features and specifications that it has, no wonder some say that it was the best phone in its range. if you are looking for a high spec but has a limited budget, this phone is the best choice!”	271	19/3/2021
A14Z9LAETO21KL	“It’s Awesome”	0	24/10/2020
A3RHT4KI3H5TVH	“Good phone with terrible battery. I am about to return it.”	5	15/1/2021
AKYDGCKCY7H9F	“Product is great has double SIM card slot; everything is great except the battery.”	12	3/01/2021
A18RGYRCEN181M	“Love it so far”	0	24/02/2021
A34VZEFXQJJ7AT	“I love this phone. I have never gotten a newer phone model but always opted for a cheaper few generation behind the phone so this is a pretty big upgrade from my galaxy s7. I have a straight talk and it worked immediately by just switching over my sim card. The phone looks sleek and	97	28/03/2021

	fits perfectly in my smaller hands (I'm a 5'0" tall woman so my hands are tiny). "		
A3GAP455S8YH0M	"The battery life sucks. I've done all of the things to optimize the battery too...dark mode, 60 instead of 120 Hz even Power saving in the settings. This battery doesn't hold a charge as long as my two-year-old note 9 so I am highly disappointed. Wishing I had gone with the Note 20."	21	24/12/2020
A3W2VF6D09B2RN	"The fingerprint reader failed after 5 days and then it unlocked with a swipe for half a day. The image of a fingerprint disappeared from the screen, but the phone unlocked with a swipe for a few hours after which a swipe only brought up a request for the password. However, the PW hadn't been recorded. "	4	19/11/2020
A2X1F0WUJJP0FC	"First time buying a phone from somewhere other than the main carriers. Upgraded from an S7 to the S20 5G. It has been an overall great phone."	780	03/04/2020

AQES9A9BCJ7CV	"I am delighted with this purchase. Everything like I said in the description. The phone is 100% original and nothing is missing. Comes with your AKG hearing aids and original charger. I'm very happy to have it."	0	20/01/2021

5.5 Conclusion

This chapter covers everything from data preparation to calculating various scores, such as the total review score and review rating scores. These data analysis methods led to the study's outcomes, which are detailed in detail in the next chapter, Chapter 6.

6 RESULTS

6.1 Introduction

This chapter aims to examine and report on the findings of the review using the data collection strategy that was chosen. In a big manner, the analysis is planned and organized and this approach is used to analyse the data in this overall research process. Section 6.2 describes the performance of fine-tuned BERT model and section 6.3, section 6.4, and section 6.5 explains the review helpfulness score, review time score and review sentiment score respectively. Section 6.6 explains the review score calculation and section 6.7 comparing the results with the existing system result of customer review evaluations. Section 6.8 concludes this chapter with a summary of this chapter.

6.2 Performance of Fine-Tuned BERT Model

After the fine-tuned process of the BERT model, it provides a very efficient and effective model for conducting an effective sentiment analysis even in large datasets. The efficiency of the finetuned BERT-Base model for this study is shown in Table 6.1.

Table 6.1 - *BERT-Base model Classification Result*

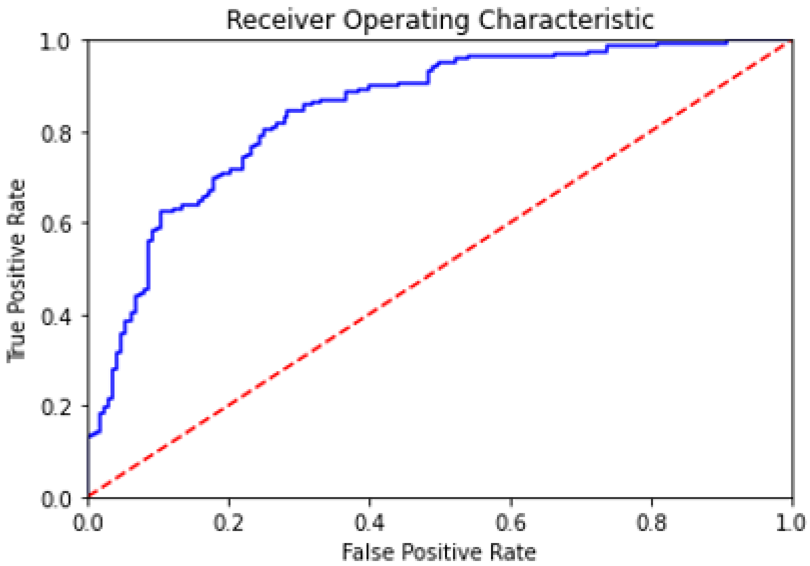
	Precision	Recall	F1 Score	Accuracy
BERT	0.88	0.89	0.88	0.88

The classification accuracy(Gu et al., 2009 and Hossin et al., 2011) is the total number of accurate predictions divided by the total number of predictions in a dataset. Through accuracy, the nature of the created arrangement is assessed depending on the level of right forecasts over absolute examples. The supplement metric of exactness is the error rate which assesses by its level of inaccurate expectations. Researchers have widely used all of these metrics to discriminate and choose the best solution (Hossin & Sulaiman, 2015).

The ratio of True Positives to all Positives is known as precision. The recall is a test of how well this model detects True Positives. Finally, the Harmonic Mean of Precision and Recall is the F1-score. Based on the findings, it may conclude that the fine-tuned BERT model achieved a high degree of accuracy, precision, and recall.

A true positive is a consequence where the positive class is accurately estimated by the model. Similarly, a true negative is a result where the negative class is accurately estimated by the model. A false positive is an outcome where the positive category is falsely estimated by the algorithm. And a false negative is a result in which the negative class is falsely predicted by the model (Google, n.d). The ROC Curve for the finetuned BERT Model is seen in Figure 6.1.

Figure 6.1 - ROC Curve



6.3 Review Helpfulness Score

As mentioned in Table 4, this study mainly considering a total of 1500 customer reviews for four manufacturers and nine different products and all of the customer reviews are also verified by amazon.com. To calculate the review helpfulness score, the researcher uses a range from 0.75 to 1 and if there is nothing as a review helpfulness vote then it will be considered as the minimum allocated score of 0.75 (refer the equation 1). The data set contains the different number of votes from 0 to 783 as review helpfulness votes. Table 6.2 illustrating sample user reviews and their corresponding review helpfulness votes.

To show the functionality of Algorithm 1, here taking Smartphone as a category and SGF20FE as a product then considering fifteen reviews randomly selected and passing the corresponding input to the algorithm for review helpfulness calculation then we get results as follows.

Table 6.2 - Random selection of reviews used to demonstrate Algorithm 1

Review ID	**No. of Helpfulness votes**
A2A039TZMZHH9Y	13

A195EZSQDW3E21	28
A3M1PLEYNDEYO8	159
A35QFQI0M46LWO	0
A1GMWTGXW682GB	4
A2FZ4Z0UFA1OR8	271
A14Z9LAETO21KL	0
A3RHT4KI3H5TVH	5
AKYDGCKCY7H9F	12
A18RGYRCEN181M	0
A34VZEFXQJJ7AT	97
A3GAP455S8YH0M	21
A3W2VF6D09B2RN	4
A2X1F0WUJJP0FC	780
AQES9A9BCJ7CV	0

Algorithm 1 following Equation 1 to calculate the review helpfulness score and the high value for total helpfulness votes received for an entity plays a vital role in the formulated equation. In this example Review ID A2FZ4Z0UFA1OR8 got a maximum number of helpfulness vote of 271 and it will be considered as the Nh_j for Equation 1. Table 6.3 provides the review helpfulness score for these randomly selected review datasets.

Table 6.3 - Review Helpfulness Score

Review ID	**No. of Helpfulness votes**	**Review Helpfulness Score**
A2A039TZMZHH9Y	13	0.750
A195EZSQDW3E21	28	0.750
A3M1PLEYNDEYO8	159	0.811
A35QFQI0M46LWO	0	0.750
A1GMWTGXW682GB	4	0.750
A2FZ4Z0UFA1OR8	271	0.893

A14Z9LAETO21KL	0	0.750
A3RHT4KI3H5TVH	5	0.750
AKYDGCKCY7H9F	12	0.750
A18RGYRCEN181M	0	0.750
A34VZEFXQJJ7AT	97	0.750
A3GAP455S8YH0M	21	0.750
A3W2VF6D09B2RN	4	0.750
A2X1F0WUJJP0FC	219	0.863
AQES9A9BCJ7CV	0	0.750

The maximum review helpfulness votes earn maximum helpfulness score and the minimum or 0 votes gets a minimum of 0.75 as helpfulness score while we following Algorithm 1.

6.4 Review Time Score

To demonstrate the review time score calculation this study will use the same set of reviews. Equation 2 along with Algorithm 2 is used to calculate the review time score based on the time that the review created. According to the researcher, the latest feedback help consumers with some more recent knowledge so this study suggests the following equation to give each analysis a time score. The researcher used 10 years as a realistic period for calculating the time score; if the period exceeded the period, the researcher used 0.8 as the review time score for each review. The researcher discovered that the rating score of 0.8 and greater than 0.8 provided successful results to the studies with the aid of several other feedback. As a result, the researcher opted to use the same score for this study as well.

The calculation for the review time score is simple and easy to understand. If the review written on a period of 10 years from the current date it will be considered for maximum review score calculation and the review score value should be between 0.8 and 1 for these kinds of user reviews. If the reviewed time exceeds the limit of ten years, then the equation assigns a minimum value of 0.8 to that particular review.

To demonstrate the functionality of Algorithm 2, consider the following scenario: taking Smartphone as a category and SGF20FE as a product, then randomly selecting ten reviews and forwarding the corresponding input to the algorithm for review helpfulness measurement, we get the following results. Table 6.4 contains the Review ID and its reviewed time and Table 6.5 illustrating the review time score output when we applying the algorithm to the selected dataset.

Table 6.4 - Review ID and Review Time to Calculate Review Time Score

Review ID	**Review Time**
A2A039TZMZHH9Y	1/12/2020
A195EZSQDW3E21	13/11/2020
A3M1PLEYNDEYO8	05/02/2021
A35QFQI0M46LWO	14/12/2020
A1GMWTGXW682GB	30/11/2020
A2FZ4Z0UFA1OR8	19/3/2021
A14Z9LAETO21KL	24/10/2020
A3RHT4KI3H5TVH	15/1/2021
AKYDGCKCY7H9F	3/01/2021
A18RGYRCEN181M	24/02/2021
A34VZEFXQJJ7AT	28/03/2021
A3GAP455S8YH0M	24/12/2020
A3W2VF6D09B2RN	19/11/2020
A2X1F0WUJJP0FC	03/04/2020
AQES9A9BCJ7CV	20/01/2021

The product SGF20FE is a newly launched product so all the reviews are within the realistic period of 10 years then we will use $1 - (Cy - Ry_{ij}) * 0.002$ as a formula. In this particular formula first calculating the difference between the current year and the year of the customer review which is created. In this particular formula first calculating the difference between the current year (Cy) and the year of the reviewed time (Ry_{ij}). The following examples discussing regarding the entire process for the review time score calculation.

Example 1:

Review ID = A2A039TZMZHH9Y

Cy= 2021; Ry_{ij}=2020

Then $Cy - Ry_{ij}$= 2021-2020=1

Then $1 - (Cy - Ry_{ij}) * 0.002$ =**0.998**

Example 2:

Cy= 2021; Ry_{ij}=2009

Then $Cy - Ry_{ij}$= 2021-2009=12

Then the Algorithm 2 assign a minimum value of **0.8** for that particular review.

As a result, the review score for considering user reviews is shown in Table6.5

Table 6.5 - Review Time Score

Review ID	**Review Time**	**Review Time Score**
A2A039TZMZHH9Y	1/12/2020	0.998
A195EZSQDW3E21	13/11/2020	0.998
A3M1PLEYNDEYO8	05/02/2021	1.000
A35QFQI0M46LWO	14/12/2020	0.998
A1GMWTGXW682GB	30/11/2020	0.998
A2FZ4Z0UFA1OR8	19/3/2021	1.000
A14Z9LAETO21KL	24/10/2020	0.998
A3RHT4KI3H5TVH	15/1/2021	1.000
AKYDGCKCY7H9F	3/01/2021	1.000
A18RGYRCEN181M	24/02/2021	1.000
A34VZEFXQJJ7AT	28/03/2021	1.000
A3GAP455S8YH0M	24/12/2020	0.998
A3W2VF6D09B2RN	19/11/2020	0.998
A2X1F0WUJJP0FC	03/04/2020	0.998
AQES9A9BCJ7CV	20/01/2021	1.000

The most recent reviews will get more score than the older ones and this score will affect when we considering the overall review score.

6.5 Review Sentiment Score

By analysing semantic associations between terms or sub-words in a text using a variety of NLP approaches, the researcher learned the BERT model to measure the feeling orientation probability of particular analysis and it helped to reach adequate outcomes. This sentiment score generation algorithm assigns a sentimental orientation score to each study. For this, the researcher devised the formula shown below and applied the max() function to the BERT Model's output. Equation 3 and Algorithm 3 is using to find the appropriate sentiment score for each user reviews. If BERT output contains negative terms than positive terms, it is considered a negative statement; otherwise, it is considered a positive statement. In this study, Positive reviews more weighted than negative reviews and each negative sentiment orientation plays an important role in the overall review score. If the review holds a negative sentiment orientation, then the BERT score will be treated as a negative value otherwise it will be considered as a positive value. The Review ID and the corresponding text (user review) used to demonstrate the sentiment orientation score is shown in Table 5.3. Table 6.6 showing both the sentiment orientation and sentiment score for the reviews.

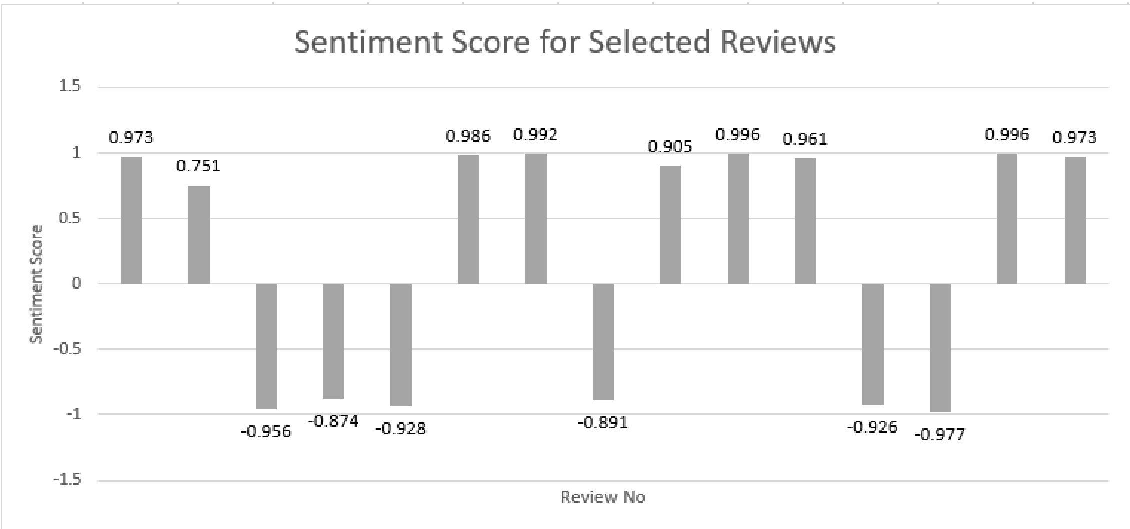
Table 6.6 - Sentiment Orientation Score for the Selected Reviews

Review No	Review ID	Sentiment Orientation	Sentiment Score
1	A2A039TZMZHH9Y	Positive	0.973
2	A195EZSQDW3E21	Positive	0.751
3	A3M1PLEYNDEYO8	Negative	-0.956
4	A35QFQI0M46LWO	Negative	-0.874
5	A1GMWTGXW682GB	Negative	-0.928
6	A2FZ4Z0UFA1OR8	Positive	0.986
7	A14Z9LAETO21KL	Positive	0.992
8	A3RHT4KI3H5TVH	Negative	-0.891
9	AKYDGCKCY7H9F	Positive	0.905
10	A18RGYRCEN181M	Positive	0.996

11	A34VZEFXQJJ7AT	Positive	0.961
12	A3GAP455S8YH0M	Negative	-0.926
13	A3W2VF6D09B2RN	Negative	-0.977
14	A2X1F0WUJJP0FC	Positive	0.996
15	AQES9A9BCJ7CV	Positive	0.973

Figure 6.1 draws the Sentiment score for different reviews based on the review number.

Figure 6.2 - Sentiment Score



A combination of review helpfulness score, review time score and review sentiment score will generate the overall review score and which will be discussed in the next section.

6.6 Review Score

Overall review score calculation is one of the most important steps in this research. Based on previous scores such as review helpfulness (HS), review time (TS), and review sentiment orientation, the overall review score redefines the review ranking (SS). The composite of the HS, TS, and SS scores is used to calculate the total ranking. The cumulative evaluation score for this report is calculated using Equation 4. To measure the individual review score in this study, simply sum up all of the scores such as helpfulness score (HS), time score (TS), and sentiment score (SS) and divide them by the total number of review score (here it is considered as three). It is just as a simple average calculation by definition. Table 6.7 showing the obtained results for Review Helpfulness Score, Review Time Score and Review Sentiment Orientation Score from previous sections and Table 6.8

representing the overall review score for each review as an output of Algorithm 4. The final review rating score is calculated using both Equation 5 and Algorithm 5 and the result is shown in Table 6.9. Table 4.2 contains the information regarding how to calculate the star rating based on the review score. So, the corresponding star rating for the customer reviews based on the guidelines from Table 4.2 is shown in Table 6.9.

Table 6.7 - Review Helpfulness Score, Review Time Score and Review Sentiment Orientation Score for the Customer Reviews

Review ID	**Review Helpfulness Score (HS)**	**Review Time Score (TS)**	**Review Sentiment Score (SS)**
A2A039TZMZHH9Y	0.750	0.998	0.973
A195EZSQDW3E21	0.750	0.998	0.751
A3M1PLEYNDEYO8	0.811	1.000	-0.956
A35QFQI0M46LWO	0.750	0.998	-0.874
A1GMWTGXW682GB	0.750	0.998	-0.928
A2FZ4Z0UFA1OR8	0.893	1.000	0.986
A14Z9LAETO21KL	0.750	0.998	0.992
A3RHT4KI3H5TVH	0.750	1.000	-0.891
AKYDGCKCY7H9F	0.750	1.000	0.905
A18RGYRCEN181M	0.750	1.000	0.996
A34VZEFXQJJ7AT	0.750	1.000	0.961
A3GAP455S8YH0M	0.750	0.998	-0.926
A3W2VF6D09B2RN	0.750	0.998	-0.977
A2X1F0WUJJP0FC	0.863	0.998	0.996
AQES9A9BCJ7CV	0.750	1.000	0.973

Table 6.8 - Review Score

Review ID	HS + TS + SS	Review Score
A2A039TZMZHH9Y	2.721	0.907
A195EZSQDW3E21	2.449	0.816
A3M1PLEYNDEYO8	0.855	0.285
A35QFQI0M46LWO	0.874	0.291
A1GMWTGXW682GB	0.820	0.273
A2FZ4Z0UFA1OR8	2.879	0.959
A14Z9LAETO21KL	2.740	0.913
A3RHT4KI3H5TVH	0.859	0.286
AKYDGCKCY7H9F	2.682	0.885
A18RGYRCEN181M	2.746	0.915
A34VZEFXQJJ7AT	2.711	0.903
A3GAP455S8YH0M	0.822	0.274
A3W2VF6D09B2RN	0.771	0.257
A2X1F0WUJJP0FC	2.857	0.952
AQES9A9BCJ7CV	2.723	0.907

Table 6.9 - Review Rating Score and Review Star Rating Based on Algorithm 5

Review ID	Review Score	Review Rating Score	Review Rating
A2A039TZMZHH9Y	0.907	90.7	★★★★★
A195EZSQDW3E21	0.816	81.6	★★★★
A3M1PLEYNDEYO8	0.285	28.5	★
A35QFQI0M46LWO	0.291	29.1	★
A1GMWTGXW682GB	0.273	27.3	★
A2FZ4Z0UFA1OR8	0.959	95.9	★★★★★

A14Z9LAETO21KL	0.913	91.3	★★★★★
A3RHT4KI3H5TVH	0.286	28.6	★
AKYDGCKCY7H9F	0.894	88.5	★★★★
A18RGYRCEN181M	0.915	91.5	★★★★★
A34VZEFXQJJ7AT	0.903	90.3	★★★★★
A3GAP455S8YH0M	0.274	27.4	★
A3W2VF6D09B2RN	0.257	25.7	★
A2X1F0WUJJP0FC	0.952	95.2	★★★★★
AQES9A9BCJ7CV	0.907	90.7	★★★★★

6.7 Comparison of Results

To understand how this algorithm redefining the automated both review and product rating than the existing system. The existing system for the majority of the eCommerce platforms is following a manual input method to rate the review or product. As discussed earlier, the existing manual recording system has a problem such as sometimes the ratings corresponding to the reviews make some contradictions to the readers. For instance, a reviewer reviewed a product as average and marked a 5-star rating will lead the buyer to some thought processes. The researcher suggesting all these algorithms - Algorithm 1 to find the review helpfulness score, Algorithm 2 is used to calculate the review time score, Algorithm 3 is used to extract the review sentiment score from the BERT model, Algorithm 4 using to observe the overall review score, and Algorithm 5 using to calculate review rating score - to discard the human interventions on review rating as well as product rating. As per the viewpoint of the researcher, he believes that reviewing the user reviews will provide more systematic output and review score and it will help future users.

In the following section, this study compares the selected reviews' manually fed review rating with the output of this experimental study. Table 6.10 comparing the results and showing how effective all these solutions. The results of the manually fed user rating and the processed user rating are clearly distinguished in Table 6.10. For instance, the consumer gave the review A3RHT4KI3H5TVH a three-star rating. Taking into account all other variables, especially the sentiment orientation, its importance was reduced to one star instead of three. According to the researcher's opinion, the one-star ranking for the analysis A3RHT4KI3H5TVH after the assessment is acceptable.

Table 6.10 - Comparison of Rating Between User Made and Obtained

Review ID	Rating (By User)	Result of this experiment		
		Review Score	Review Rating Score	Rating
A2A039TZMZHH9Y	★★★★★	0.907	90.7	★★★★★
A195EZSQDW3E21	★★★	0.816	81.6	★★★★
A3M1PLEYNDEYO8	★★	0.285	28.5	★
A35QFQI0M46LWO	★	0.291	29.1	★
A1GMWTGXW682GB	★	0.273	27.3	★
A2FZ4Z0UFA1OR8	★★★★★	0.959	95.9	★★★★★
A14Z9LAETO21KL	★★★★	0.913	91.3	★★★★★
A3RHT4KI3H5TVH	★★★	0.286	28.6	★
AKYDGCKCY7H9F	★★★★★	0.885	88.5	★★★★
A18RGYRCEN181M	★★★★★	0.915	91.5	★★★★★
A34VZEFXQJJ7AT	★★★★	0.903	90.3	★★★★★
A3GAP455S8YH0M	★	0.274	27.4	★
A3W2VF6D09B2RN	★	0.257	25.7	★
A2X1F0WUJJP0FC	★★★	0.952	95.2	★★★★★
AQES9A9BCJ7CV	★★★★	0.907	90.7	★★★★★

On the one hand, the product SGF20FE receives 3.5 stars as a product rating based solely on randomly chosen ratings and its manual rating, and on the other hand, we earn 3.2 stars as a product rating based on the experiment. As any of the datasets used in this experiment are compared, they all display certain differences from the manually labelled user ranking, so, the researcher claims that the experiment result has more accuracy than the labelled one. Figure 6.2 and Table 6.11 showing the result comparison for all the datasets which are used in this study with the experiment output.

Figure 6.3 - User Review Rating vs Samthripthi Model Rating

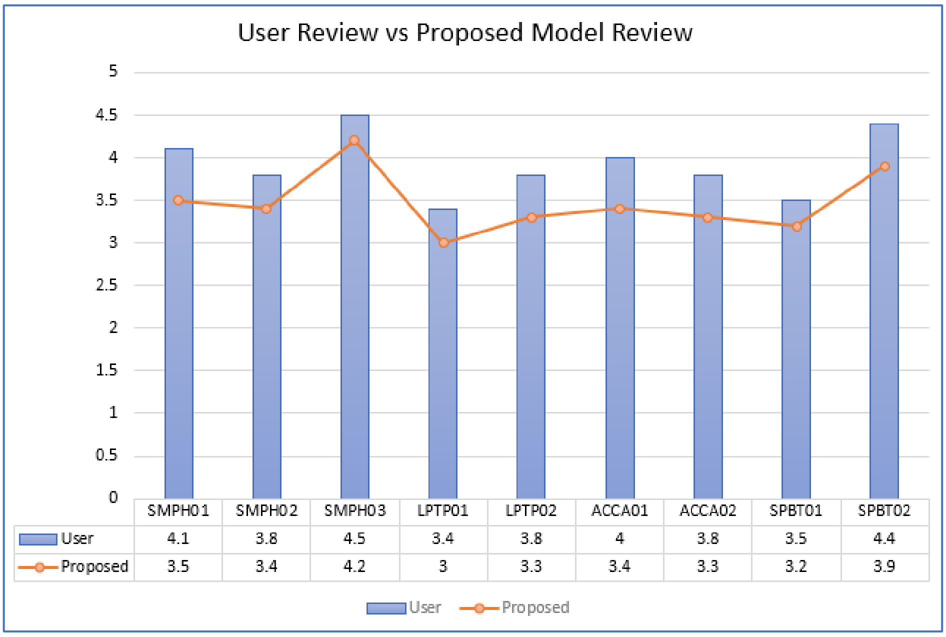


Table 6.11 - Comparison of Results

Category	Product	Product Rating (user)	Product Rating (experiment)	Difference in Rating (to user marked Rating)	Accuracy
Smart Phones	SMPH01	4.1	3.5	0.6	85.53%
	SMPH02	3.8	3.4	0.4	89.74%
	SMPH03	4.5	4.2	0.3	93.14%
Laptops	LPTP01	3.4	3.0	0.4	88.32%
	LPTP02	3.8	3.5	0.3	91.89%
Action Camera	ACCA01	4.0	3.4	0.6	85.02%
	ACCA02	3.8	3.3	0.5	86.48%

Speakers	SPBT01	3.5	3.2	0.3	90.86%
	SPBT02	4.4	3.9	0.5	88.36%

When we equate the proposed system to the current system, the researcher finds that it has an accuracy level of more than 85%. Many of the datasets are processed with an accuracy ratio of more than 90%. As a result, the researcher concludes that it is a new suitable paradigm for evaluating consumer reviews and ranking products.

6.8 Conclusion

This chapter described how each algorithm obtains the results in each step and explores the results of the whole experiment with many datasets. When comparing the proposed scheme to the existing system, the researcher discovers that it is more than 80% accurate. As a result, the researcher concludes that it is a new viable paradigm for analysing and rating customer feedback. However, this analysis method has some limitations. As a result, the next chapter discusses the study's shortcomings, prospective research proposals, and conclusions.

7 DISCUSSION

7.1 Introduction

The discussion chapter examines the context, importance and significance of the findings. It will concentrate on demonstrating and analysing the study's findings, demonstrate how they contribute to the literature review and the questions of science and argue that the final result is supported. The researcher will share his interpretation of the results and any new perceptions that have arisen as a result of this study in this discussion chapter.

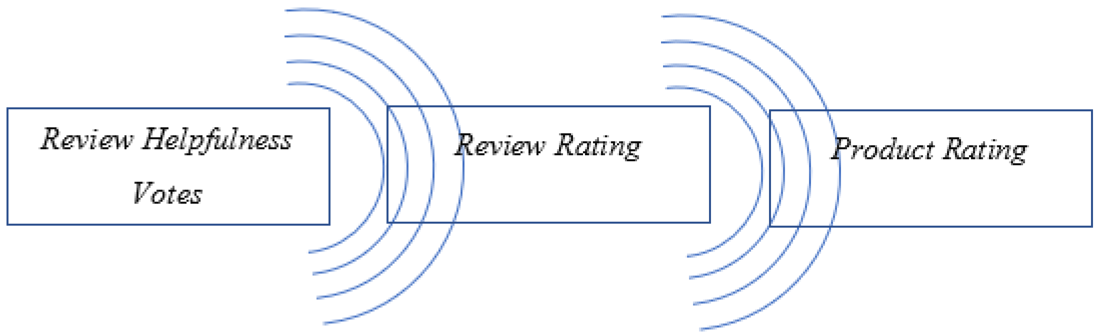
Section 6.2, section 6.3, and section 6.4 discuss this research's core element and is about the review helpfulness score, review time score, and review sentiment score. Discussion on review score and review rating is in section 6.5. Finally, section 6.6 concludes this chapter with an overall idea of this chapter.

7.2 Review Helpfulness Score

Online review created by users is now an essential component of eCommerce platform, especially the Business to Consumer (B2C) eCommerce sector. In eCommerce, online customer reviews mainly fulfil two purposes. One mission is to support potential customers in evaluating goods and services before making purchases. The other will help the customers get more details about a product (Park et al., 2007; Clemons et al., 2006; Chen & Xie, 2008; Mudambi & Schuff, 2010). To improve shoppers' buying experience, Amazon.com has revolutionised numerous online review features. As a result, the shopper will vote on his usefulness by having to click on the yes or no icon of the review material after the latest online review has been updated and read by a Registered Shopper. The survey's overall amount of yes votes will be shown as a helpful indication somewhere at the end of the consumer reviews. Based on added helpfulness votes, amazon.com can dynamically classify the most helpful responses and position them at the top of the comment section using its algorithms. This essential feature seems to be quite beneficial because when the number of reviews increases, the users cannot read and cover even its small percentage (Ghose & Ipeirotis, 2011; Mudambi & Schuff, 2010; Korfiatis et al., 2012).

In any case, the most helpful reviews have an influential capacity on users because the user trusts that as an experience of the product (Wan et al., 2007; Wan & Nakayama, 2014). So, in this research, the researcher considering review helpfulness as an essential factor while playing with customer reviews. As mentioned in section 4.3, the researcher would like to give a range of 0.75 to 1 as the review helpfulness score because the researcher believes that sometimes there are only a few helpfulness votes for entities. As a result, high review helpfulness remarks earn a high score for helpfulness and the poor helpfulness remarks receive a low level of scores as shown in Table 6.3. Thus, Hypothesis1 is supported. Figure 7.2 depicts how the review helpfulness votes influence the overall product rating based on this newly proposed method. A detailed analysis of the results clearly explaining how the review helpfulness score affect the review rating. When the number of votes for the helpfulness of a review increases, you can find a change in the overall review score.

Figure 7.1 - Influence of Review Helpfulness Votes on Product Rating



On the other hand, Bias in online reviews has also been discovered in many studies. Similar or associated prejudices can even be influenced in the review helpfulness votes (Li & Hitt, 2008; Kapoor & Piramuthu, 2009; Cui et al., 2012; Purnawirawan et al., 2012). However, this study following a criterion and considering only the reviews marked as a verified purchase, then the credibility for the review is high compared to the other unmarked reviews.

7.3 Review Time Score

As discussed earlier, only a few studies addressed the factor review time on a review evaluation and product rating. So, this study may be a revolutionary thought regarding the review time factor. As per the researcher, the age of the review will affect the user preferences and the relevance of the comment. In this study, the researcher believes that when a system is planned to automated a product rating, it should be relevant and more accurate than the previous studies and final results. The reviewed time factor will also show how seasonality impacts online feedback and the product's market presence.

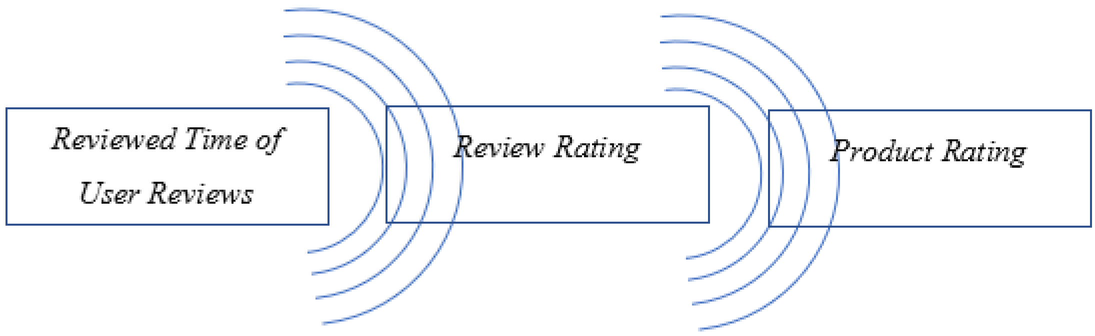
One of the most heated doubt on review time is what timeframe is realistic. In this study, the period is set at ten years, a long time for technology products relative to mechanical products. The researcher used 10 years as an essential term, which can be revised after more thorough future trials and consumer acceptance. When we calculated the review score based on a ten-year realistic period, there was almost the same time score for all the reviews.

According to the results, the review time affects the overall product rating for all period on review rating and the product rating. For instance, take a look at the following Example 1, the helpfulness score, review time score and review sentiment score for the review is 0.75, 1 and 0.996, respectively. As a result, we get 0.915 as a review score and a five-star rating for that particular review, and it is the same as the user-defined manual star rating.

Table 7.1 - Importance of Reviewed Time

	Review ID	**HS**	**TS**	**SS**	**RS**	**Rating**
Example 1	A18RGYRCEN181M	0.750	1.000	0.996	0.915	5 Stars

Figure 7.2 - Influence of Reviewed Time on Product Rating



7.4 Review Sentiment Score

Different types of ratings paint evaluations with a big brush, leaving out certain nuances about how the consumer feels. A four-star or five-star rating will seem amazing, and you might expect that the material that follows would be positive, but this is not always the case. This is where sentiment analysis proves its worth. A five-star score, for example, might list elements of the experience that the reviewer disliked despite enjoying it overall. eCommerce providers are using sentiment analysis to capture the feeling behind a consumer rating and begin to consider what factors led to or detracted from a favourable experience. For example, a five-star rating could also state that the items are suitable for throwing. Sentiment analysis enables one to decipher the subtleties of consumer feedback and trace the source of a problem (or a good experience). So, the review sentiment is a significant factor in customer review evaluation. Many studies (Haque et al., 2018; Shrestha & Nasoz, 2019; Tamara & Milicevic, 2018; Sygkounas et al., 2016; Srujan et al., 2018) indicating the importance of sentiment evaluation in online customer reviews and developed different review rating algorithms instead of the currently using one.

There are many other available methods (as mentioned in Chapter 2), but BERT is considered one of the most efficient models for sentiment analysis. From the literature review and the mentioned studies, the review score calculation without a review sentiment evaluation is not fruitful. So, this research explores all the BERT model features to attain an effective sentiment score from each review. Apart from that, sentiment polarity is one of the critical factors in this model to obtain an automated product rating. Finally, this research does not doubt the relationship between the sentiment polarity of the online consumer review and the overall product rating score. Figure 7.4 depicts how the review sentiment polarity influences the overall product rating based on this newly proposed method.

Figure 7.3 - Influence of Review Sentiment Polarity on Product Rating

7.5 Review Score and Review Rating

Past research on evaluating customer feedback based on mining customer opinions articulated in natural language has primarily concentrated on analysing the semantic and sentiment relationships of online reviews. On the other hand, Consumer ratings provide a wealth of additional material that may be included during the product rating process such as review submission time and review helpfulness votes. Unfortunately, no researches have yet combined review time, helpfulness votes, and emotion polarity to generate a single numerical product ranking score.

As discussed in Chapter 2 and Chapter 5, all the factors such as review sentiment, review time, and review helpfulness redefine the user review rating system with this new model to automate the review rating and the product rating from the customer review itself. In addition, in the comparison between the manually marked star rating and the automated star rating, the automated rating is matching more than eighty-five percentage with the manual entry.

7.6 Evidence of Research Achievements

To produce an appropriate review rating score, the system considers review helpfulness, review time, and review sentiment orientation likelihood. This new type of product ranking provides consumers with enough knowledge about the desired product to make an informed decision. The framework can be used by all websites that allow online customers to upload their reviews, exchange their ratings and vote for valuable reviews. The scheme can also be extended to different domains. The model is beneficial for providing online consumers with ample knowledge throughout their selection in e-commerce and saving time and energy in reading a thousand online review text to make an exact decision. The differences between prior online review evaluation schemes and our current product rating system as seen in Table 7.2.

Table 7.2 - Review Attributes used by Other Researchers.

Researchers	Review Attributes			
	Semantic	Sentiment	Review Helpfulness	Review Time
Yan et al., 2017	✓	✗	✗	✗
Sindhu et al., 2017	✓	✓	✗	✗
Anshuman et al., 2017	✓	✓	✗	✗
Benlahbib & Nfaoui, 2019	✓	✓	✗	✗
This Study	✓	✓	✓	✓

The researcher uses Table7.3 to present the research concisely. Columns in the table describe the primary research question, sub research questions, hypotheses for each sub research question, research evidence, and deliverables or artefacts. The page numbers and section numbers that support the specific research sub-question with its hypothesis are handled in the research evidence section. This research's deliverables are included under the deliverables and artefacts section of Table 7.3.

Table 7.3- Evidence of Research Achievements

RQ	SRQ	Hypothesis	Research Evidence		Deliverables / Artefacts
			Page No & Section	**Description**	
RQ1	RQ1.1	H1- Approved	Page 23 Section 2.9	Helps to understand the importance of review helpfulness.	1. Systematic Literature Review 2. Problem Definition and Problem Specification 3. Equations and Algorithms to evaluate user reviews based on this proposed model 'Samthripthi'.
			Page 53 Section 4.3	Describing how to calculate the review helpfulness score.	
			Page 98 Section 7.3	Describing the influence of review helpfulness in overall review rating.	
			Page 94 Table 6.11	Depicts the review rating comparison with user marked reviews against the output of the proposed model 'Samthripthi'.	
	RQ1.2	H2- Approved	Page 24 Section 2.10	Describing the importance of review time in user reviews.	
			Page 55 Section 4.4	Discussing how to calculate the review helpfulness score.	
			Page 99 Section 7.4	Describing the influence of review helpfulness in overall review rating.	

			Page 94 Table 6.11	Depicts the review rating comparison with user marked reviews against the output of the proposed model called 'Samthripthi'.	
	RQ1.3	H3-Approved	Page 24 Section 2.11	Describing the importance of review sentiment orientation of the user reviews for forecasting its effects to an automated review rating or product rating system.	
			Page 58 Section 4.5	Discussing how to calculate the review sentiment orientation score using finetuned BERT model.	
			Page 100 Section 7.5	Describing the influence of review sentiment in overall review rating.	
			Page 94 Table 6.11	Depicts the review rating comparison with user marked reviews against the output of the proposed model 'Samthripthi'.	

7.7 Limitations

Firstly, the chance for biased reviews is not negligible. Many studies (discussed in Section 6.3) already found that online customer reviews are biased in some cases. Sometimes the reviews are generated to promote some wrong information (negative and positive) to the users. The verified purchase badge helps overcome this problem to some extent but needs some more studies to identify the biased reviews.

Second, as previously said, the practical timeframe in this analysis is set at ten years, which is a long time for technological products compared to mechanical products. The researcher used a 10-year time frame as a starting point, which can be adjusted based on subsequent experiments and market adoption results. When we measured the review score over a ten-year reasonable timeframe, it found that all of the reviews had almost identical time scores. In addition, in the case of sentiment analysis, the online user can mislead the sentiment by applying idioms instead of actual words. BERT model considering it to some extent but need a foolproof finetuned solution to overcome this. In addition to this, the data is collected for a particular time and not for many different intervals. So, the question is open to its trend-following capacity.

In addition, the lack of user experience and expert views are some of the most considerable limitations of this study. To better understand this proposed model 'Samthripthi', the researcher can test it with consumers and seek feedback from experts. This type of feedback would aid in the improvement of the current product rating model's performance.

Finally, the whole experiment is conducted only a small dataset for selected products. To identify its credibility and accuracy, we need to make this experiment in a largescale dataset. Due to the time and fund this experiment using WebCrawler which is open source or availed at less cost and it has its limitations such as the number of reviews can fetch from amazon.com.

7.8 Conclusion

The findings and their interpretations are discussed in this chapter. For more precise findings, this study believes that this analysis should have been extended by adjusting the sample size and discussing the shortcomings discussed in section 6.3. This segment also has suggestions for overcoming those limits. The proposals for future research and the conclusion of the report are presented in the following chapter.

8 CONCLUSION

8.1 Introduction

The satisfied client uses the user review option better to clarify the product features and functionality to new customers. Dissatisfied customers have chosen it as a simple way to convey their frustration, so eWOM, especially online user feedback, is one type of speech by both satisfied and dissatisfied ones. Based on the modified DSR Model, the researcher consulted previous studies and articles on product ranking based on customer review results to assess the research questions. Then, based on the published literature, he determined that there isn't a solution that takes review sentiment, review period validity, and review helpfulness into account. When the factors mentioned above are taken into account, the researcher concludes that the study about a new solution for product ranking is essential. This paper proposes a product rating system that mines consumer and user feedback articulated in natural language to create credibility for different products. Review helpfulness, review period, review sentiment polarity, and review ranking are included in the scheme. In previous studies, the researchers considered review sentiment analysis an effective technique to evaluate customer reviews rather than consider the other factors. As discussed in section 2.6 there have several methods to assess the sentiment of the user reviews. Still, due to the bidirectional feature - the entire text passage is taken into account to explain the meaning of each word - of the BERT model and a comparison between BERT and other models proves that the BERT has more accuracy on sentiment analysis compared to Vanilla CNN, Vanilla LSTM and Vanilla BiLSTM. Thus, this study following finetuned BERT model for better results. Finally, the researcher proposes a solution named 'Sampthrithi' (സംതൃപ്തി), which is regarded as a practical solution by this research by combining the BERT model with the algorithms stated in chapter 4. But this study has some limitations that are discussed in section 6.7. So, it should need some more future studies listed as in section 7.2. Section 7.3 concluding this chapter as well as this study.

8.2 Future Work

This study primarily focused on eWOM or online consumer feedback, especially on the Amazon.com eCommerce website, and discovered that attributes such as helpfulness votes, checked time, and sentiment polarity of the review would influence the overall product rating when measured using the algorithms or equations proposed in this study. There are some other opportunities to expand this thesis or consider it a new possible research project, which is discussed further below.

- Future studies should exploit additional features such as user credibility (prolific reviewers) and online behaviour.
- Future research can attempt to identify and delete false and irrelevant reviews through the use of a filtering phase, minimising processing time and enhance the device performance at the same time, so only genuine and usable reviews would be considered.
- Future research can consider reputation visualisation by displaying the top 5 positive and top 5 negative reviews, which will directly help the user to decide without investing much more time in the reviews.
- The impact of eWOM through social media is one of the demanding future research options regarding the user review evaluation to explore the available opportunities for brand reputation.
- Another research is needed to determine the significance and duration of its review-time factor. The researcher hopes that subsequent studies will have an accurate and fool proof practical time frame where the proposed model ‘Samthripthi’ is used to rate products.
- Finally, future investigations can include the tackling mechanism for idioms which is currently misleading the sentiment analysis.

8.3 Concluding Remarks

User review analysis has been a must for e-commerce in this day and age. Today, it is critical to creating an unbiased product rating based on examining consumer feelings derived from their textual review. This suggested rating system aids in gaining more accurate and impartial perspectives into customer opinions on product quality and brand tastes; moreover, this research collects consumer feedback to determine the degree of consumer satisfaction. The overall reflection for the researcher can be concluded as follows. The researcher acquired new expertise during the study. When doing research, the researcher learned new skills and concepts, such as concentrating on the research subject and articulating the topic. The report's data was compiled after reading and analysing numerous separate documents. As a part of this research, the researcher extracted the dataset from amazon.com and applied the proposed model ‘Samthripthi’ then he got a favourable result which accepts his all hypothesis. The researcher believes that the proposed system will shed light on the research based on customer review evaluation and product rating, especially the attributes considered in this research and the usage of the BERT model throughout this research. However, this study has its drawbacks. Further studies should be undertaken to achieve more extensive and more reliable outcomes in the future.

References

Agarwal, B., & Mittal, N. (2015). Semantic orientation-based approach for sentiment analysis. *Socio-Affective Computing*, 77-88. https://doi.org/10.1007/978-3-319-25343-5_6

Ahluwalia, R. (2002). How prevalent is the negativity effect in consumer environments? *Journal of Consumer Research*, *29*(2), 270-279. https://doi.org/10.1086/341576

Aken, J. E. (2004). Management research based on the paradigm of the design sciences: The quest for field-tested and grounded technological rules. *Journal of Management Studies*, *41*(2), 219-246. https://doi.org/10.1111/j.1467-6486.2004.00430.x

Alturki, A., Gable, G. G., & Bandara, W. (2011). A design science research roadmap. *Service-Oriented Perspectives in Design Science Research*, 107-123. https://doi.org/10.1007/978-3-642-20633-7_8

Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. *Academy of Management Review*, *36*(2), 247-271. https://doi.org/10.5465/amr.2011.59330882

Amazon.com. (n.d.). https://www.amazon.com/gp/help/customer/display.html?nodeId=GQUXAMY73JFRVJHE#:~:text=Amazon%20calculates%20a%20product's%20star,the%20authenticity%20of%20the%20feedback

Anshuman, Rao, S., & Kakkar, M. (2017). A rating approach based on sentiment analysis. *2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence*. https://doi.org/10.1109/confluence.2017.7943213

Archak, N., Ghose, A., & Ipeirotis, P. G. (2011). Deriving the pricing power of product features by mining consumer reviews. *Management Science*, *57*(8), 1485-1509. https://doi.org/10.1287/mnsc.1110.1370

Augustyniak, Ł., Szymański, P., Kajdanowicz, T., & Tuligłowicz, W. (2015). Comprehensive study on lexicon-based ensemble classification sentiment analysis. *Entropy*, *18*(1), 4. https://doi.org/10.3390/e18010004

Babić Rosario, A., De Valck, K., & Sotgiu, F. (2019). Conceptualizing the electronic word-of-mouth process: What we know and need to know about eWOM creation, exposure, and evaluation. *Journal of the Academy of Marketing Science*, *48*(3), 422-448. https://doi.org/10.1007/s11747-019-00706-1

Bart, Y. (2017). Product seeding: Word-of-Mouth effects for and beyond the focal product. *GfK Marketing Intelligence Review*, *9*(2), 24-29. https://doi.org/10.1515/gfkmir-2017-0014

Baskerville, R. L., Kaul, M., & Storey, V. C. (2015). Genres of inquiry in design-science research: Justification and evaluation of knowledge production. *MIS Quarterly*, *39*(3), 541-564. https://doi.org/10.25300/misq/2015/39.3.02

Becker, J., Vom Brocke, J., Heddier, M., & Seidel, S. (2015). In search of information systems (Grand) challenges. *Business & Information Systems Engineering*, *57*(6), 377-390. https://doi.org/10.1007/s12599-015-0394-0

Benlahbib, A., & Nfaoui, E. H. (2019). A hybrid approach for generating reputation based on opinions fusion and sentiment analysis. *Journal of Organizational Computing and Electronic Commerce*, *30*(1), 9-27. https://doi.org/10.1080/10919392.2019.1654350

Benlahbib, A., & Nfaoui, E. H. (2019). An unsupervised approach for reputation generation. *Procedia Computer Science*, *148*, 80-86. https://doi.org/10.1016/j.procs.2019.01.011

Benlahbib, A., & Nfaoui, E. H. (2020). Aggregating customer review attributes for online reputation generation. *IEEE Access*, *8*, 96550-96564. https://doi.org/10.1109/access.2020.2996805

Bi, J., Liu, Y., & Fan, Z. (2019). Representing sentiment analysis results of online reviews using interval type-2 fuzzy numbers and its application to product ranking. *Information Sciences*, *504*, 293-307. https://doi.org/10.1016/j.ins.2019.07.025

Bickart, B., & Schindler, R. M. (2001). Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, *15*(3), 31-40. https://doi.org/10.1002/dir.1014

Bigné, E., Caplliure, E., & Miquel, M. (2016). EWOM on travel agency selection: Specialized versus private label. *Psychology & Marketing*, *33*(12), 1046-1053. https://doi.org/10.1002/mar.20938

Breck, E., & Cardie, C. (2017). Opinion mining and sentiment analysis. *The Oxford Handbook of Computational Linguistics 2nd edition*. https://doi.org/10.1093/oxfordhb/9780199573691.013.43

Breck, E., & Cardie, C. (2017). undefined. *The Oxford Handbook of Computational Linguistics 2nd edition*. https://doi.org/10.1093/oxfordhb/9780199573691.013.43

Burghardt, M. (2011). The predictive power of retail investor sentiment. *Retail Investor Sentiment and Behavior*, 119-142. https://doi.org/10.1007/978-3-8349-6170-9_6

Büschken, J., & Allenby, G. M. (2016). Sentence-based text analysis for customer reviews. *Marketing Science*, *35*(6), 953-975. https://doi.org/10.1287/mksc.2016.0993

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, *28*(2), 15-21. https://doi.org/10.1109/mis.2013.30

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems*, *50*(2), 511-521. https://doi.org/10.1016/j.dss.2010.11.009

Cardie, C. (2017). Keeping apace with progress in natural language processing. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. https://doi.org/10.1145/3018661.3022745

Chanji, A., & Bhatia, R. (2016). Various techniques for sentiment analysis of products review – A review. *International Journal of Computer Applications*, *144*(8), 29-32. https://doi.org/10.5120/ijca2016910422

Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-Mouth as a new element of marketing communication mix. *Management Science*, *54*(3), 477-491. https://doi.org/10.1287/mnsc.1070.0810

Chen, Y., & Yao, S. (2017). Sequential search with refinement: Model and application with click-stream data. *Management Science*, *63*(12), 4345-4365. https://doi.org/10.1287/mnsc.2016.2557

Cheung, C. M., Lee, M. K., & Rabjohn, N. (2008). The impact of electronic word-of-mouth. *Internet Research*, *18*(3), 229-247. https://doi.org/10.1108/10662240810883290

Clemons, E., Guodong Gao, & Hitt, L. (2006). When online reviews meet Hyperdifferentiation: A study of craft beer industry. *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*. https://doi.org/10.1109/hicss.2006.534

Cormen, T. H. (2009). *Introduction to algorithms*. MIT Press.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE.

Cui, G., Lui, H., & Guo, X. (2012). The effect of online consumer reviews on new product sales. *International Journal of Electronic Commerce*, *17*(1), 39-58. https://doi.org/10.2753/jec1086-4415170102

De los Santos, B. I., & Koulayev, S. (2011). Optimizing click-through in online rankings for partially Anonymous consumers. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2163365

De, S. R. (2017). Sentiment analysis on product purchase through E commerce. *International Journal of Scientific Research and Management*. https://doi.org/10.18535/ijsrm/v5i6.13

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, *49*(10), 1407-1424. https://doi.org/10.1287/mnsc.49.10.1407.17308

Devlin, J., & Chang, M. W. (2018). *Open sourcing BERT: State-of-the-Art pre-training for natural language processing*. Google AI Blog. https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. https://arxiv.org/pdf/1810.04805.pdf

Dewi, C. K., & Mohaidin, Z. (2016). Motivations of online opinion seeking and its effect on the online purchase intention. *Advanced Science Letters*, *22*(12), 4073-4076. https://doi.org/10.1166/asl.2016.8144

Dixit, S., & Santhosh, K. R. (2016). Collaborative analysis of customer feedbacks using rapid miner. *International Journal of Computer Applications*, *142*(2), 29-36. https://doi.org/10.5120/ijca2016909695

Dominik, H. (2017). Sentiment analysis. *Data Science – Analytics and Applications*, 111-112. https://doi.org/10.1007/978-3-658-19287-7_17

Dresch, A., Lacerda, D. P., & Antunes, J. A. (2014). Design science research. *Design Science Research*, 67-102. https://doi.org/10.1007/978-3-319-07374-3_4

D'souza, S. R., & Sonawane, K. (2019). Sentiment analysis based on multiple reviews by using machine learning approaches. *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*. https://doi.org/10.1109/iccmc.2019.8819813

Dwivedi, S., Patel, H., & Sharma, S. (2019). Movie reviews classification using sentiment analysis. *Indian Journal of Science and Technology*, *12*(41), 1-6. https://doi.org/10.17485/ijst/2019/v12i41/145554

Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, *2*(1). https://doi.org/10.1186/s40537-015-0015-2

Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, *73*(3), 421-435. https://doi.org/10.1037/0021-9010.73.3.421

Fernández-Gavilanes, M., Álvarez-López, T., Juncal-Martínez, J., Costa-Montenegro, E., & Javier González-Castaño, F. (2016). Unsupervised method for sentiment analysis in online texts. *Expert Systems with Applications*, *58*, 57-75. https://doi.org/10.1016/j.eswa.2016.03.031

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*(3), 291-313. https://doi.org/10.1287/isre.1080.0193

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction*. Allyn & Bacon.

Gadekallu, T., Soni, A., Sarkar, D., & Kuruva, L. (2019). undefined. *Advances in Business Information Systems and Analytics*, 77-90. https://doi.org/10.4018/978-1-5225-4999-4.ch006

Ghose, A., & Ipeirotis, P. G. (2011). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Transactions on Knowledge and Data Engineering*, *23*(10), 1498-1512. https://doi.org/10.1109/tkde.2010.188

Ghose, A., Ipeirotis, P. G., & Li, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, *60*(7), 1632-1654. https://doi.org/10.1287/mnsc.2013.1828

Godes, D. (2013). Product policy in markets with word-of-Mouth communication. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2275617

Godes, D., & Mayzlin, D. (2004). Using online conversations to study word-of-Mouth communication. *Marketing Science*, *23*(4), 545-560. https://doi.org/10.1287/mksc.1040.0071

Goldsmith, R. E., & Horowitz, D. (2006). Measuring motivations for online opinion seeking. *Journal of Interactive Advertising*, *6*(2), 2-14. https://doi.org/10.1080/15252019.2006.10722114

Google. (n.d.). *TP vs. FP rate at different classification thresholds*. https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc

Gu, Q., Zhu, L., & Cai, Z. (2009). Evaluation measures of the classification performance of Imbalanced data sets. *Communications in Computer and Information Science*, 461-471. https://doi.org/10.1007/978-3-642-04962-0_53

Guellil, I., & Boukhalfa, K. (2015). Social big data mining: A survey focused on opinion mining and sentiments analysis. *2015 12th International Symposium on Programming and Systems (ISPS)*. https://doi.org/10.1109/isps.2015.7244976

Gundecha, P., & Liu, H. (2012). Mining social media: A brief introduction. *2012 TutORials in Operations Research*, 1-17. https://doi.org/10.1287/educ.1120.0105

Guo, C., Du, Z., & Kou, X. (2017). Mining online customer reviews for products aspect-based ranking. *Communications in Computer and Information Science*, 150-161. https://doi.org/10.1007/978-981-10-6989-5_13

Haddara, M., Hsieh, J., Fagerstrøm, A., Eriksson, N., & Sigurðsson, V. (2019). Exploring customer online reviews for new product development: The case of identifying reinforcers in the cosmetic industry. *Managerial and Decision Economics*, *41*(2), 250-273. https://doi.org/10.1002/mde.3078

Haque, T. U., Saber, N. N., & Shah, F. M. (2018). Sentiment analysis on large scale Amazon product reviews. *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*. https://doi.org/10.1109/icird.2018.8376299

Hemalatha, S., & Ramathmika, R. (2019). Sentiment analysis of yelp reviews by machine learning. *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*. https://doi.org/10.1109/iccs45141.2019.9065812

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing*, *18*(1), 38-52. https://doi.org/10.1002/dir.10073

Hevner, A., & Chatterjee, S. (2010). Design science research in information systems. *Integrated Series in Information Systems*, 9-22. https://doi.org/10.1007/978-1-4419-5653-8_2

Hlee, S., Lee, H., & Koo, C. (2018). undefined. *Sustainability*, *10*(4), 1141. https://doi.org/10.3390/su10041141

Horowitz. (1997). *Fundamentals of computer algorithms*. Galgotia Publications.

Hossin, M., Sulaiman, M., Mustapha, A., Mustapha, N., & Rahmat, R. (2011). A hybrid evaluation metric for optimizing classifier. *2011 3rd Conference on Data Mining and Optimization (DMO)*. https://doi.org/10.1109/dmo.2011.5976522

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04*. https://doi.org/10.1145/1014052.1014073

Huang, A. H., Chen, K., Yen, D. C., & Tran, T. P. (2015). A study of factors that contribute to online review helpfulness. *Computers in Human Behavior*, *48*, 17-27. https://doi.org/10.1016/j.chb.2015.01.010

huggingface.co. (n.d.). Hugging Face – The AI community building the future. https://huggingface.co/transformers/#:~:text=%F0%9F%A4%97%20Transformers%20(formerly%20known%20as,in%20100%2B%20languages%20and%20deep

Ismagilova, E., Dwivedi, Y. K., Slade, E., & Williams, M. D. (2017). Impact of eWOM. *Electronic Word of Mouth (eWOM) in the Marketing Context*, 73-96. https://doi.org/10.1007/978-3-319-52459-7_6

Jana, E., & Uma, V. (2019). Aspect extraction and sentiment analysis for e-Commerce product reviews. *Emerging Trends in Computing and Expert Technology*, 780-791. https://doi.org/10.1007/978-3-030-32150-5_79

Jeziorski, P., & Segal, I. (2015). What makes them click: Empirical analysis of consumer demand for search advertising. *American Economic Journal: Microeconomics*, *7*(3), 24-53. https://doi.org/10.1257/mic.20100119

Jianqiang, Z., & Xiaolin, G. (2017). Comparison research on text pre-processing methods on Twitter sentiment analysis. *IEEE Access*, *5*, 2870-2879. https://doi.org/10.1109/access.2017.2672677

Johnson, R., & Zhang, T. (2015). Effective use of word order for text categorization with Convolutional neural networks. *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. https://doi.org/10.3115/v1/n15-1011

Kanayama, H., & Nasukawa, T. (2006). Fully automatic lexicon expansion for domain-oriented sentiment analysis. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing - EMNLP '06*. https://doi.org/10.3115/1610075.1610125

Kanayama, H., Nasukawa, T., & Watanabe, H. (2011). A sentiment analysis method using tree transfer model. *Transactions of the Japanese Society for Artificial Intelligence*, *26*, 273-283. https://doi.org/10.1527/tjsai.26.273

Kapoor, G., & Piramuthu, S. (2009). Sequential bias in online product reviews. *Journal of Organizational Computing and Electronic Commerce*, *19*(2), 85-95. https://doi.org/10.1080/10919390902820764

Karamibekr, M., & Ghorbani, A. A. (2013). A structure for opinion in social domains. *2013 International Conference on Social Computing*. https://doi.org/10.1109/socialcom.2013.44

Karimi, S., & Wang, F. (2017). Online review helpfulness: Impact of reviewer profile image. *Decision Support Systems*, *96*, 39-48. https://doi.org/10.1016/j.dss.2017.02.001

Kaushik, K., Mishra, R., Rana, N. P., & Dwivedi, Y. K. (2018). Exploring reviews and review sequences on e-Commerce platform: A study of helpful reviews on Amazon.in. *Journal of Retailing and Consumer Services*, *45*, 21-32. https://doi.org/10.1016/j.jretconser.2018.08.002

Kawakami, T., Kishiya, K., & Parry, M. E. (2012). Personal word of mouth, virtual word of mouth, and innovation use. *Journal of Product Innovation Management*, *30*(1), 17-30. https://doi.org/10.1111/j.1540-5885.2012.00983.x

Khattak, A., Paracha, W. T., Asghar, M. Z., Jillani, N., Younis, U., Saddozai, F. K., & Hameed, I. A. (2020). Fine-grained sentiment analysis for measuring customer satisfaction using an extended set of fuzzy linguistic hedges. *International Journal of Computational Intelligence Systems*, *13*(1), 744. https://doi.org/10.2991/ijcis.d.200513.001

Kim, J. B., Albuquerque, P., & Bronnenberg, B. J. (2009). Online demand under limited consumer search. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1340267

Knuth, D. E. (1997). *The art of computer programming: Fundamental algorithms*. Addison-Wesley Professional.

Korfiatis, N., García-Bariocanal, E., & Sánchez-Alonso, S. (2012). Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications*, *11*(3), 205-217. https://doi.org/10.1016/j.elerap.2011.10.003

Kudeshia, C., & Kumar, A. (2017). Social eWOM: Does it affect the brand attitude and purchase intention of brands? *Management Research Review*, *40*(3), 310-330. https://doi.org/10.1108/mrr-07-2015-0161

Kumar, A., & Abhirami, S. (2015). An experimental study of feature extraction techniques in opinion mining. *International Journal on Soft Computing, Artificial Intelligence and Applications*, *4*(1), 15-21. https://doi.org/10.5121/ijscai.2015.4102

Kumar, A., & Sebastian, T. M. (2012). Sentiment analysis: A perspective on its past, present and future. *International Journal of Intelligent Systems and Applications*, *4*(10), 1-14. https://doi.org/10.5815/ijisa.2012.10.01

Labrecque, L. I., Esche, J. V., Mathwick, C., Novak, T. P., & Hofacker, C. F. (2013). Consumer Power: Evolution in the Digital Age. *Journal of Interactive Marketing*. https://doi.org/10.1016/j.intmar.2013.09.002

Leavy, P. (2017). *undefined*. Guilford Publications.

Lee, M. K., Cheung, C. M., Lim, K. H., & Ling Sia, C. (2006). Understanding customer knowledge sharing in web-based discussion boards. *Internet Research*, *16*(3), 289-303. https://doi.org/10.1108/10662240610673709

Lei Wu, Hoi, S. C., & Nenghai Yu. (2010). Semantics-preserving bag-of-Words models and applications. *IEEE Transactions on Image Processing*, *19*(7), 1908-1920. https://doi.org/10.1109/tip.2010.2045169

Li, X., & Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information Systems Research*, *19*(4), 456-474. https://doi.org/10.1287/isre.1070.0154

Li, X., Wu, C., & Mai, F. (2019). The effect of online reviews on product sales: A joint sentiment-topic analysis. *Information & Management*, *56*(2), 172-184. https://doi.org/10.1016/j.im.2018.04.007

Liang, P., & Dai, B. (2013). Opinion mining on social media data. *2013 IEEE 14th International Conference on Mobile Data Management*. https://doi.org/10.1109/mdm.2013.73

Liu, B. (2011). Opinion mining and sentiment analysis. *Web Data Mining*, 459-526. https://doi.org/10.1007/978-3-642-19460-3_11

Liu, B. (2015). *Opinions, sentiment, and emotion in text*. Cambridge University Press.

Liu, C., & Liao, K. (2019). Efficiently predicting frequent patterns over uncertain data streams. *Procedia Computer Science*, *160*, 15-22. https://doi.org/10.1016/j.procs.2019.09.438

Liu, Q. (2015). Online WOM effects and product type: Evidence from Tmall. *Proceedings of the 2015 Joint International Mechanical, Electronic and Information Technology Conference*. https://doi.org/10.2991/jimet-15.2015.101

Liu, Y., Bi, J., & Fan, Z. (2017). Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory. *Information Fusion*, *36*, 149-161. https://doi.org/10.1016/j.inffus.2016.11.012

Liu, Y., & Zhu, X. (2018). Short text sentiment classification based on feature extension and ensemble classifier. https://doi.org/10.1063/1.5039023

Loos, P., Nebel, W., Marx Gómez, J., Hasan, H., Watson, R. T., Vom Brocke, J., Seidel, S., & Recker, J. (2011). Green IT: A matter of business and information systems engineering? *Business & Information Systems Engineering*, *3*(4), 245-252. https://doi.org/10.1007/s12599-011-0165-5

M, H., & M.N, S. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, *5*(2), 01-11. https://doi.org/10.5121/ijdkp.2015.5201

Ma, Q., & Zhang, W. (2015). Public mood and consumption choices: Evidence from sales of Sony cameras on Taobao. *PLOS ONE*, *10*(4), e0123129. https://doi.org/10.1371/journal.pone.0123129

Maharani, W. (2013). Microblogging sentiment analysis with lexical based and machine learning approaches. *2013 International Conference of Information and Communication Technology (ICoICT)*. https://doi.org/10.1109/icoict.2013.6574616

Malik, M. S. (2020). Predicting users' review helpfulness: The role of significant review and reviewer characteristics. *Soft Computing*, *24*(18), 13913-13928. https://doi.org/10.1007/s00500-020-04767-1

Mate, C. (2015). Survey on product review sentiment analysis with aspect ranking. *International Journal of Science and Research (IJSR)*, *4*(12), 1783-1786. https://doi.org/10.21275/v4i12.nov152367

Mehto, A., & Indras, K. (2016). Data mining through sentiment analysis: Lexicon based sentiment analysis model using aspect catalogue. *2016 Symposium on Colossal Data Analysis and Networking (CDAN)*. https://doi.org/10.1109/cdan.2016.7570901

Metzger, M. J. (2007). Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, *58*(13), 2078-2091. https://doi.org/10.1002/asi.20672

Miao, Q., Li, Q., & Zeng, D. (2010). Fine-grained opinion mining by integrating multiple review sources. *Journal of the American Society for Information Science and Technology*, *61*(11), 2288-2299. https://doi.org/10.1002/asi.21400

Min, H., Yun, J., & Geum, Y. (2018). Analyzing dynamic change in customer requirements: An approach using review-based Kano analysis. *Sustainability*, *10*(3), 746. https://doi.org/10.3390/su10030746

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, *339*(jul21 1), b2535-b2535. https://doi.org/10.1136/bmj.b2535

Mudambi, & Schuff. (2010). Research note: What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185. https://doi.org/10.2307/20721420

Offermann, P., Levina, O., Schönherr, M., & Bub, U. (2009). Outline of a design science research process. *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology - DESRIST '09*. https://doi.org/10.1145/1555619.1555629

Ogihara, M. (2018). Online and recursive algorithms. *Fundamentals of Java Programming*, 485-505. https://doi.org/10.1007/978-3-319-89491-1_19

O'Connell, M., Griffith, R., & Smith, K. (2014). The importance of product reformulation versus consumer choice in improving diet quality. https://doi.org/10.1920/wp.ifs.2014.1415

Ozer Sarı, F., & Asman Alikilic, Ö. (2016). How ready are the Turkish hospitality and travel organizations for E-complaint handling? *AJIT-e Online Academic Journal of Information Technology*, 75-94. https://doi.org/10.5824/1309-1581.2016.3.005.x

Pan, Y., & Zhang, J. Q. (2011). Born unequal: A study of the helpfulness of user-generated product reviews. *Journal of Retailing*, *87*(4), 598-612. https://doi.org/10.1016/j.jretai.2011.05.002

Pandey, A. (2016). Sentiment analysis through text Mining-A review. *International Journal Of Engineering And Computer Science*. https://doi.org/10.18535/ijecs/v5i10.52

Park, D., Lee, J., & Han, I. (2007). The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce*, *11*(4), 125-148. https://doi.org/10.2753/jec1086-4415110405

Park, J. (2020). Framework for sentiment-driven evaluation of customer satisfaction with cosmetics brands. *IEEE Access*, *8*, 98526-98538. https://doi.org/10.1109/access.2020.2997522

Pavlou, P. A., & Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Information Systems Research*, *17*(4), 392-414. https://doi.org/10.1287/isre.1060.0106

Pecar, S. (2018). Towards opinion summarization of customer reviews. *Proceedings of ACL 2018, Student Research Workshop*. https://doi.org/10.18653/v1/p18-3001

Peng, Y., Kou, G., & Li, J. (2014). A fuzzy PROMETHEE approach for mining customer reviews in Chinese. *Arabian Journal for Science and Engineering*, *39*(6), 5245-5252. https://doi.org/10.1007/s13369-014-1033-7

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. https://doi.org/10.3115/v1/d14-1162

Pownall, C. (2015). The furious customer. *Managing Online Reputation*, 97-105. https://doi.org/10.1057/9781137382306_8

Pruzan, P. (2016). Scientific method and the design of research. *Research Methodology*, 175-205. https://doi.org/10.1007/978-3-319-27167-5_7

Purnawirawan, N., De Pelsmacker, P., & Dens, N. (2012). Balance and sequence in online reviews: How perceived usefulness affects attitudes and intentions. *Journal of Interactive Marketing*, *26*(4), 244-255. https://doi.org/10.1016/j.intmar.2012.04.002

Puschmann, C., & Powell, A. (2018). Turning words into consumer preferences: How sentiment analysis is framed in research and the news media. *Social Media + Society*, *4*(3), 205630511879772. https://doi.org/10.1177/2056305118797724

Racherla, P., & King, R. A. (2012). What we know and don't know about online word-of-Mouth: A systematic review and synthesis of the literature. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2187040

Ratan, S., Anand, T., & Ratan, J. (2019). Formulation of research question – Stepwise approach. *Journal of Indian Association of Pediatric Surgeons*, *24*(1), 15. https://doi.org/10.4103/jiaps.jiaps_76_18

Reubens, R. (2019). To craft, by design, for sustainability. *Holistic Sustainability Through Craft-Design Collaboration*, 49-64. https://doi.org/10.4324/9781351065665-4

Rossetti, M., Stella, F., Cao, L., & Zanker, M. (2014). Analysing user reviews in tourism with topic models. *Information and Communication Technologies in Tourism 2015*, 47-58. https://doi.org/10.1007/978-3-319-14343-9_4

Saif, H., He, Y., Fernandez, M., & Alani, H. (2016). Contextual semantics for sentiment analysis of Twitter. *Information Processing & Management*, *52*(1), 5-19. https://doi.org/10.1016/j.ipm.2015.01.005

Saldaña, Z. W. (2018). Sentiment analysis for exploratory data analysis. *The Programming Historian*, (7). https://doi.org/10.46430/phen0079

Santhi Priya, P., & Venkateswara Rao, T. (2018). Sentiment analysis of review data of a product using Python. *International Journal of Engineering & Technology*, *7*(3.12), 674. https://doi.org/10.14419/ijet.v7i3.12.16452

Sasikala, S. P., & Sheela, M. I. (2020). Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS. https://doi.org/10.21203/rs.2.19872/v2

Savyanavar, A., & Mehta, B. (2016). Multi-document summarization using TF-IDF algorithm. *International Journal Of Engineering And Computer Science*. https://doi.org/10.18535/ijecs/v5i4.40

Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, *80*(2), 159-169. https://doi.org/10.1016/j.jretai.2004.04.001

Seunghee Im. (2017). The relationship between consumers' WOM motivations and the valence of WOM on movie. *Journal of Product Research*, *35*(6), 125-132. https://doi.org/10.36345/kacst.2017.35.6.013

Sharma, J., & Rather, R. A. (2015). Understanding the customer experience: An exploratory study of a category hotels. *International Journal on Customer Relations*, *3*(2). https://doi.org/10.21863/ijcr/2015.3.2.010

Shrestha, N., & Nasoz, F. (2019). Deep learning sentiment analysis of Amazon.Com reviews and ratings. *International Journal on Soft Computing, Artificial Intelligence and Applications*, *8*(1), 01-15. https://doi.org/10.5121/ijscai.2019.8101

Sijoria, C., Sengupta, S., & Mukherjee, S. (2017). What makes eWOM viral? *International Journal of Internet Marketing and Advertising*, *11*(4), 287. https://doi.org/10.1504/ijima.2017.10007886

Sindhu, C., Vyas, D. V., & Pradyoth, K. (2017). Sentiment analysis based product rating using textual reviews. *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)*. https://doi.org/10.1109/iceca.2017.8212762

Sivasakthi, K. (2020). As an analysis of different classification technique using sentiment analysis of product review. *Journal of Advanced Research in Dynamical and Control Systems*, *12*(SP3), 622-628. https://doi.org/10.5373/jardcs/v12sp3/20201300

Snijders, C., & Matzat, U. (2019). Online reputation systems. *The Oxford Handbook of Gossip and Reputation*, 478-495. https://doi.org/10.1093/oxfordhb/9780190494087.013.25

Soper, D. S., & Turel, O. (2012). Who are we? Mining institutional identities using N-grams. *2012 45th Hawaii International Conference on System Sciences*. https://doi.org/10.1109/hicss.2012.642

Springer US. (2017). TF–IDF. *Encyclopedia of Machine Learning and Data Mining*, 1274-1274. https://doi.org/10.1007/978-1-4899-7687-1_832

Srinivasan, S. (2017). Marketing applications using big data. *Studies in Big Data*, 487-502. https://doi.org/10.1007/978-3-319-53817-4_18

Srujan, K. S., Nikhil, S. S., Raghav Rao, H., Karthik, K., Harish, B., & Keerthi Kumar, H. (2018). Classification of Amazon book reviews based on sentiment analysis. *Advances in Intelligent Systems and Computing*, 401-411. https://doi.org/10.1007/978-981-10-7512-4_40

Stringam, B. B., & Gerdes, J. (2010). An analysis of word-of-Mouse ratings and guest comments of online hotel distribution sites. *Journal of Hospitality Marketing & Management*, *19*(7), 773-796. https://doi.org/10.1080/19368623.2010.508009

Sudha, M. K. (2020). Social media sentiment analysis for opinion mining. *International Journal of Psychosocial Rehabilitation*, *24*(5), 3672-3679. https://doi.org/10.37200/ijpr/v24i5/pr202075

Sygkounas, E., Rizzo, G., & Troncy, R. (2016). Sentiment polarity detection from Amazon reviews: An experimental study. *Semantic Web Challenges*, 108-120. https://doi.org/10.1007/978-3-319-46565-4_8

Tamara, K., & Milicevic, N. (2018). Comparing sentiment analysis and document representation methods of Amazon reviews. *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)*. https://doi.org/10.1109/sisy.2018.8524814

Tsvetkov, Y., Faruqui, M., & Dyer, C. (2016). Correlation-based intrinsic evaluation of word vector representations. *Proceedings of the 1st Workshop on Evaluating Vector-Space Representations for NLP*. https://doi.org/10.18653/v1/w16-2520

Tuckman, B. W., & Harper, B. E. (2012). *Conducting educational research*. Rowman & Littlefield Publishers.

Turney, P. D. (2001). Thumbs up or thumbs down? *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*. https://doi.org/10.3115/1073083.1073153

Tuveri, F., & Angioni, M. (2013). An opinion mining model for generic domains. *Distributed Systems and Applications of Information Filtering and Retrieval*, 51-64. https://doi.org/10.1007/978-3-642-40621-8_3

Ursu, R. M. (2018). The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science, 37*(4), 530-552. https://doi.org/10.1287/mksc.2017.1072

Verberne, S., Boves, L., Oostdijk, N., & Coppen, P. (2010). What is not in the bag of words for Why-QA? *Computational Linguistics*, *36*(2), 229-245. https://doi.org/10.1162/coli.09-032-r1-08-034

Vo, T. T., Xiao, X., & Ho, S. Y. (2017). How does corporate social responsibility engagement influence word of mouth on Twitter? Evidence from the airline industry. *Journal of Business Ethics*, *157*(2), 525-542. https://doi.org/10.1007/s10551-017-3679-z

Vom Brocke, J. (2007). Design principles for reference modelling. *Advances in Database Research*, 269-296. https://doi.org/10.4018/978-1-60566-278-7.ch014

Vyas, V., & Uma, V. (2019). Approaches to sentiment analysis on product reviews. *Advances in Business Information Systems and Analytics*, 15-30. https://doi.org/10.4018/978-1-5225-4999-4.ch002

Waila, P., Marisha, Singh, V. K., & Singh, M. K. (2012). Evaluating machine learning and unsupervised semantic orientation approaches for sentiment analysis of textual reviews. *2012 IEEE International Conference on Computational Intelligence and Computing Research*. https://doi.org/10.1109/iccic.2012.6510235

Wan, Y., Menon, S., & Ramaprasad, A. (2007). A classification of product comparison agents. *Communications of the ACM*, *50*(8), 65-71. https://doi.org/10.1145/1278201.1278208

Wan, Y., & Nakayama, M. (2014). The reliability of online review helpfulness. *Journal of Electronic Commerce Research*, *15*(3), 179-189. https://web.csulb.edu/journals/jecr/issues/20143/Paper3.pdf

Wang, B., Zhu, W., & Chen, L. (2008). Improving the Amazon review system by exploiting the credibility and time-decay of public reviews. *2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. https://doi.org/10.1109/wiiat.2008.30

Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the web. *Journal of the American Society for Information Science and Technology*, *53*(2), 134-144. https://doi.org/10.1002/asi.10016

Wieringa, R. J. (2014). *Design science methodology for information systems and software engineering*. Springer.

Wu, H. C., Luk, R. W., Wong, K. F., & Kwok, K. L. (2008). Interpreting TF-IDF term weights as making relevance decisions. *ACM Transactions on Information Systems*, *26*(3), 1-37. https://doi.org/10.1145/1361684.1361686

Wu, J. (2017). Review popularity and review helpfulness: A model for user review effectiveness. *Decision Support Systems*, *97*, 92-103. https://doi.org/10.1016/j.dss.2017.03.008

Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, *44*, 120-130. https://doi.org/10.1016/j.ijhm.2014.10.013

Yan, S., Kuperberg, G. R., & Jaeger, T. F. (2017). Prediction (or not) during language processing. A commentary on Nieuwland et al. (2017) and Delong et al. (2005). https://doi.org/10.1101/143750

Yan, Z., Jing, X., & Pedrycz, W. (2017). Fusing and mining opinions for reputation generation. *Information Fusion*, *36*, 172-184. https://doi.org/10.1016/j.inffus.2016.11.011

Yin, D., Bond, S. D., & Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly*, *38*(2), 539-560. https://doi.org/10.25300/misq/2014/38.2.10

Yuanlin Chen, Yueting Chai, Yi Liu, & Yang Xu. (2015). Analysis of review helpfulness based on consumer perspective. *Tsinghua Science and Technology*, *20*(3), 293-305. https://doi.org/10.1109/tst.2015.7128942

Zacharia, G., Moukas, A., & Maes, P. (2000). Collaborative reputation mechanisms for electronic marketplaces. *Decision Support Systems*, *29*(4), 371-388. https://doi.org/10.1016/s0167-9236(00)00084-1

Zehra, S., Wasi, S., Jami, I., Nazir, A., Khan, A., & Waheed, N. (2017). Ontology-based sentiment analysis model for recommendation systems. *Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. https://doi.org/10.5220/0006491101550160

Zhang, K., Cheng, Y., Liao, W., & Choudhary, A. (2012). Mining millions of reviews. *Proceedings of the 13th International Conference on Electronic Commerce - ICEC '11*. https://doi.org/10.1145/2378104.2378116

Zhang, Y., Jin, R., & Zhou, Z. (2010). Understanding bag-of-words model: A statistical framework. *International Journal of Machine Learning and Cybernetics*, *1*(1-4), 43-52. https://doi.org/10.1007/s13042-010-0001-0

Zheng-Jun Zha, Jianxing Yu, Jinhui Tang, Meng Wang, & Tat-Seng Chua. (2014). Product aspect ranking and its applications. *IEEE Transactions on Knowledge and Data Engineering*, *26*(5), 1211-1224. https://doi.org/10.1109/tkde.2013.136

Zhu, M., & Lai, S. (2009). A study about the WOM influence on tourism destination choice. *2009 International Conference on Electronic Commerce and Business Intelligence*. https://doi.org/10.1109/ecbi.2009.104